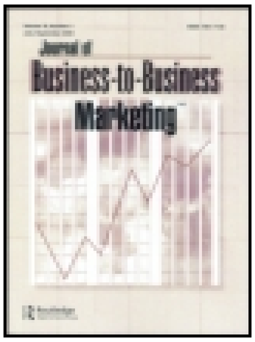

Journal of Business-to-Business Marketing

ISSN: 1051-712X (Print) 1547-0628 (Online) Journal homepage: https://www.tandfonline.com/loi/wbbm20

# A Case Study of Fintech Industry: A Two-Stage Clustering Analysis for Customer Segmentation in the B2B Setting

Alireza Sheikh, Tohid Ghanbarpour & Davoud Gholamiangonabadi

JOURNAL OF BUSINESS-TO-BUSINESS MARKETING
https://doi.org/10.1080/1051712X.2019.1603420

Routledge
Taylor & Francis Group

PRACTITIONER NOTE

Check for updates

# A Case Study of Fintech Industry: A Two-Stage Clustering Analysis for Customer Segmentation in the B2B Setting

Alireza Sheikh[a], Tohid Ghanbarpour[b], and Davoud Gholamiangonabadi[b]

[a]Department of Management, Science and Technology, Amirkabir University of Technology, Tehran, Iran; [b]Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran

**ABSTRACT**

**Purpose:** This study proposes a new approach considering two-stage clustering and LRFMP model (Length, Recency, Frequency, Monetary and Periodicity) simultaneously for customer segmentation and behavior analysis among the Iranian Fintech companies.
**Methodology/Approach**: In this study, the K-means clustering algorithm and LRFMP model are combined in the customer segmentation process. After initial clustering, for a better understanding of valuable customers, additional clustering is implemented in segments that needed further investigation. This approach contributes to the better interpretation of different customer segments. Finally, customer segments, consisting of 23524 B2B customers, are analyzed based on their characteristics and appropriate strategies are recommended accordingly.
**Findings**: The first stage clustering result shows that customers are best segmented into four groups which are named as "loyal and valuable customers", "old and churned customers", "young and churned customers" and "young and valuable customers". The first and fourth segments are clustered again and the final 11 groups of customers are determined.
**Research implications:** This study extends the RFM model and customer segmentation literature by applying the LRFMP model, which is recently used in retail sector for the first time, in the B2B setting in general and the Fintech sector in particular. It demonstrates, once again, the applicability of customer relationship and segmentation methods in the B2B setting. Also, the two-stage clustering approach adopted in this paper not only increases the precision of customer segmentation, but also subsequently enhances the efficiency of marketing strategies towards them.
**Practical implications**: This study provides a systematic and practical approach for researchers and practitioners for segmentation, interpretation and targeting of customers especially in the B2B setting and the Fintech industry. It helps managers to make effective marketing strategies and enhances customer relationship and marketing intelligence as well.
**Originality/Value/Contribution of the Paper:** This study contributes to the customer segmentation and customer relationship management literature building on and applying both the RFM and LRFMP models in the B2B setting by proposing a new approach based on a two-stage clustering method which assists to a more in-depth understanding of customer behavior.

**KEYWORDS**
Customer Segmentation; RFM Model; B2B; Fintech Industry; Cluster Analysis

## Introduction

Organizations today are seeking multifarious strategies and ways to expand their reach, efficiency, and loyalty in their customer bases. Effective communication with the customers is reported to have a positive impact on the company's profits. The field of customer relationship management is grown, advanced and matured over the past decade. Customer relationship management is mainly theorized based on four dimensions (Ngai, Xiu and Chan 2009). These four dimensions are customer identification, customer attraction, customer retention, and customer development.

In addition to identifying valuable and loyal business customers, classification of customers can lead to a better understanding of the behavior and preferences of different customer groups and help businesses to devise efficient strategies for each customer group (Dibb 1998). The use of systematic data analysis for the identification of and communication with the customers has become an important criterion in customer relationship management (Li, Dai, and Tseng 2011). As a result, many companies deploy data-driven methods to find the customers' characteristics in order to develop their market strategies

**CONTACT** Alireza Sheikh a.sheikh@aut.ac.ir
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/wbbm.

accordingly (Kim, Suh, and Hwang 2003). In this regard, the RFM model is a behavior-based model that is used to analyze customer behavior. The RFM model is a common tool for developing marketing strategies as well. This model, have also used to improve the marketing plans and increase feedback rates (Sohrabi and Khanlari 2007), in determining customers to be targeted and 'managed' (Colombo and Jiang 1999). One of the most efficient approaches for analyzing customers' behavior is utilizing data mining techniques which not only can help with the segmentation of the customers but also for developing appropriate policies for customer relationship management.

Although many studies have used variations of RFM models and data mining techniques to assess the value of customers and their ranking in various industries, such studies in the financial technologies industry and in particular, the online payment system providers, have remained sparse so far. In view of this gap, the LRFMP model proposed by (Peker, Kocyigit, and Eren 2017), which was used to investigate the customers in the grocery retail industry, is adopted in this study. The behavior of the business customers of a payment processor company in Iran's financial technology industry have examined. This research is the first of its kind carried out at the B2B level using LRFMP variables and via a two-stage clustering method for the segmentation of customers.

## Literature considerations

In the customer segmentation literature, the main focus has mainly been on the RFM variables thus far. The RFM model is widely used in various fields, including financial and nonprofit organizations (Hsieh 2004; Sohrabi and Khanlari 2007), government agencies (King 2007), online industries (Li, Lin, and Lai 2010), communication industries (Li, Shue, and Lee 2008) and the marketing industry in particular (Jonker, Piersma, and Potharst 2006; Spring, Leeflang, and Wansbeek 1999).

In this regard and in the B2B setting, the weighted RFM model was used to assess the loyalty of the customers of Iran's SAPCO Company in 2009. This model was implemented using the k-means algorithm, and its optimal cluster number was determined by the DBI index. The results of this study showed that weighting of the components of the RFM model could improve the results and increase the accuracy of the formation of clusters (Hosseini, Maleki, and Gholamian 2010).

In another study, the customers of an FMCG company in Egypt were divided into eight clusters using the LRFM model. This study was conducted using the k-means ++ algorithm (Kandeil, Saad, and Youssef 2014). Another variant of the LRFMP model, in which there is an additional P factor as an indicator of customer profitability, was used in a two-stage study to predict customer churn in the logistics industry. By identifying business-critical customers, the LRFMP model boosts the performance of the customer churn prediction model (Chen, Hu, and Hsieh 2015).

As stated above, most of the studies have been conducted using different variants of the RFM model and in the B2C business setting. In one study, the customers of a UK retailer shop were segmented according to the RFM model using the k-means algorithm and by applying a decision tree, the main characteristics of the customers in each cluster were determined. In this study, 3726 customers were divided into five clusters (Chen, Sain, and Guo 2012).

Another study in Taiwan employed the k-means algorithm and the RFM model to divide as many as 675 customers of a retail business into six clusters. The results of this study showed that a large part of the company's resources is spent on two clusters that contribute to the most of the firm's profits. Two other clusters of the customers need more resources to be allocated in order to become profitable (Wu, Chang, and Lo 2009). In another study on the cosmetics industry, the RFM model and a modified model which used the additional factor of the number of purchased items were compared with each other. The results ultimately did not show any difference. This study was conducted using the k-means algorithm as well (Khajvand et al. 2011).

In some businesses such as grocery stores, the periodicity of customer returns is an important factor in analyzing the customers' behavior. Hence, the LRFMP model was developed for this study where the P factor implies the periodicity of the customer return. A study in Turkey used this model by applying the k-means algorithm to analyze the customers of a grocery store (Peker, Kocyigit, and Eren 2017).

The review of the relevant literature shows that there is a dearth of research not only in the B2B setting in general, but also the sector of payment processor companies in particular which are heavily reliant on customer segmentation and relationship management. In most studies to date, only one clustering level is used whereas, in this study, two layers of clustering have been used to have a better understanding of the customers and to choose the most appropriate strategies. Therefore, in this practitioner note, the customers of an intermediary payment service provider company in Iran are investigated using the LRFMP model and they are segmented into clusters. Regarding this explanation, the scientific contribution of this practitioner note is twofold; first, the development of thematic literature and the choice of appropriate model for the B2B business model, especially in payment service provider companies. Second, a multi-level clustering method has been used simultaneously for better understanding of the customers' behavior.

## Methodology

This section elucidates the methods used in this research. First, the pre-processing of data is explained. Then, by applying the obtained data, the segmentation of customers using the clustering technique is carried out. The results of the clustering technique will be considered in view of the Davies-Bouldin index for cluster validation.

### *Data preprocessing*

There are different normalization algorithms, such as Min–Max normalization, Z-score normalization, and sigmoid normalization. In many studies in the field of consumer behavior analysis, Z-score method has been used for pre-processing of the data (Peker, Kocyigit, and Eren 2017). In this practitioner note, too, the Z-score method is used for data standardization. The Z-Score method transforms the data into a distribution with the mean of 0 and a standard deviation of 1. The operators *mean (: j)* and *STD (: j)* denote the arithmetic mean and standard deviation operators for feature $j$, respectively:

$$data(i,j) = \frac{data(i,j) - mean(:,j)}{std(:j)} \tag{1}$$

Where $data(i,j)$ shows the value of record $i$-th and attribute $j$-th.

### *K-means algorithm*

Clustering is the practice of grouping the data into similar groups. A cluster is a set of data that have the closest similarity and proximity to each other within that cluster, as well as the most non-similarity with the data inside other clusters (Khajvand et al. 2011). The K-Means algorithm is one of the most well-known clustering algorithms which is commonly used in customer clustering and RFM analyses. The steps for implementing this algorithm as deployed in this research are as follows:

I. First, K points are selected as the points of the cluster centers.

II. Each data is assigned to the cluster whose center has the smallest distance to that data.

III. After all the data are assigned to each of the clusters, the center of the new cluster is calculated based on the average of the points belonging to it.

IV. The 2nd and 3rd steps are repeated until there are no changes in the cluster centers.

### *Davies–bouldin index*

This index is an internal factor and evaluates clustering validity (Davies and Bouldin 1979). The algorithm and procedure for calculating this index is as follows. Assume that $C_i$ denotes the $i$-th cluster of our data and $X_j$ denotes the $j$-th data sample with $n$ attributes for the $i$-th cluster. Define $S_i$ as:

$$S_i = \left( \frac{1}{T_i} \sum_{j=1}^{T_i} \left| X_j - A_i \right|^q \right)^{\frac{1}{q}} \tag{2}$$

Where $A_i$ is the center of the $i$-th cluster and $T_i$ denotes the number of members inside that cluster. Now, $M_{i,j}$ is defined as the distance between the two clusters $i$ and $j$ based on the following formula:

$$M_{i,\,j} = \| \mathrm{A\_i} - \mathrm{A\_j} \|_p = \left( \sum_{k=1}^{n} \left| a_{k,\,i} - a_{k,\,j} \right|^p \right)^{\frac{1}{p}} \tag{3}$$

Where $a_{k,i}$ is the $k$-th attribute of the $A_i$ element (center of the $i$-th cluster).

Furthermore, $R_{i,j}$ is defined as an indicator of the quality or fairness of the clustering between the $i$-th and $j$-th clusters. The Davies-Bouldin index is defined as a ratio of $S_i$ and $M_{i,\,j}$, which has to satisfy the following conditions.

$$R_{i,j} \geq 0 \tag{4}$$

$$R_{i,\,j} = R_{j,\,i} \tag{5}$$

if $S_j \geq S_k$ and $M_{i,j} = M_{i,k}$ then $R_{i,j} > R_{i,k}$ (6)

if $S_j = S_k$ and $M_{i,j} \leq M_{i,k}$ then $R_{i,j} > R_{i,k}$ (7)

Regarding these conditions, $R_{i,j}$ can be defined as

$$R_{i,j} = \frac{S_i + S_j}{M_{i,\,j}} \tag{8}$$

Lower values of this index is an indicator of greater separation between the clusters and greater concentration within the clusters.

Now let:

$$D_i \equiv \max_{j:i \neq j} R_{i,\,j} \tag{9}$$

If the total number of clusters is$N$, the Davies-Bouldin index is defined as follows.

$$DB \equiv \frac{1}{N} \sum_{i=1}^{N} D_i \tag{10}$$

Figure 1 schematically illustrates the process of conducting this research. As shown in this figure, first the data are pre-processed and then, the initial clustering is carried out. After that, the required clusters are re-clustered again for further investigation. By doing so, one can obtain more detailed information and insight about the behavior of the customers in these clusters. Lastly, the necessary analysis is presented for both types of clustering.

## Data

The data used in this study includes the information on the value, the number and the date of the first and last transactions for 23524 customers during a 45-month period ending on 13 June 2018. This information is collected from the users who had at least one transaction in the past 12 months, which has been extracted from the operational database of one of the online payment processor companies in the Iranian Fintech industry. As explicated in the review of the literature, the following variables are defined for this study as listed in Table 1.

The statistical information of the data is presented in Table 2.

In Table 2, the mean values, standard deviations, maximum and minimum values for L, R, F, M, and P variables are expressed. On average, the duration of customers' relationship with the company is less than one year. The average number of purchases per each customer is about 885, and the average transaction value rate is about 31216674 Rials (40000 Rials = 1 $). In this database, there were also some customers who were considered as valid data, but they only had transactions at the last day of data collection, and some other customers had financial transactions on this day as well. A list of businesses that used the services of this intermediary payment company is provided in Table 3. The carrier distribution of these businesses is also shown in Figure 2.

As is evident in Figure 2, most of the related businesses were in the area of *educational tools* with more than 3500 cases. In terms of the largest number of cases in each group, the next big group was the *courses* with nearly 1500. On the other hand, the business groups of *software, books*, and *flash-games* had the smallest numbers.

## Results

Regarding the methodology presented in Section 3, first, the customer data were normalized. Then, the k-means method was used to cluster the customers. The results of the Davies-Bouldin index for different clustering techniques are shown in Table 4.

In order to cluster the customers, the k-means clustering method has been used. For each number of clusters, the model was executed five times to get the best clustering results as presented in Table 4. Among these variants of clustering, the one with four clusters has had the lowest value of the Davies-Bouldin index. Therefore, it can be concluded that this clustering (with four clusters) is more compact

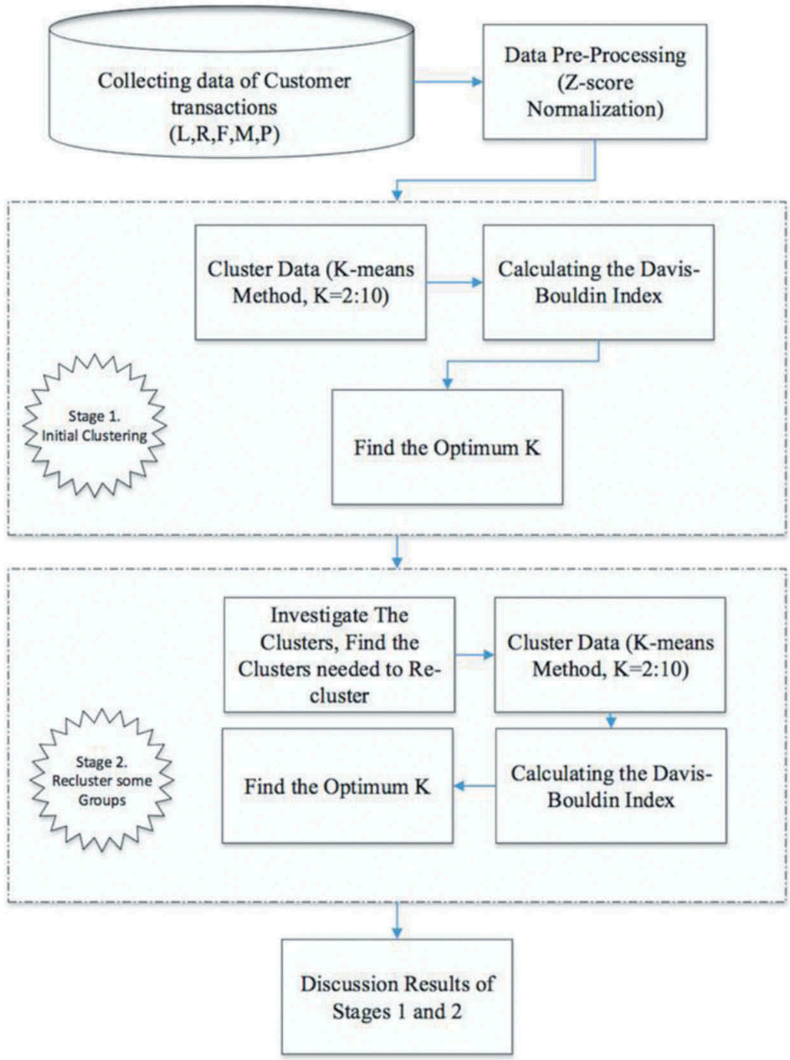


**Figure 1.** Proposed Approach.

**Table 1.** Definition of parameters.

| Parameter | Definition | Adapted parameter |
|---|---|---|
| Length | Length of relationship between customer and the business | Interval between the first transaction and the last transaction |
| Recency | Last purchase date in a particular period | Interval between the last transaction and the end of the period |
| Frequency | Number of purchases in a particular period | Number of transactions which occurred in the period |
| Monetary | Value of purchases in a particular period | The aggregate of transactions value during the period |
| Periodicity | Regularity of customers | The average interval between days which transaction(s) occurred |
| Monetary: IRR | | Length, Recency, Periodicity: Day(s) |

and; thus, it has embedded customers that are more resembling one another.

**Table 2.** Description of raw data.

| | L | R | F | M | P |
|---|---|---|---|---|---|
| Mean | 254.81 | 107.89 | 884.73 | 31216674 | 22.06 |
| STD | 291.77 | 110.74 | 16232.63 | 692382625.8 | 35.83 |
| Max | 1368 | 364 | 1612553 | 81751159398 | 568 |
| Min | 0 | 0 | 1 | 100 | 0 |

As can be seen in Table 5, some of the values in this table are distinct from the other data. For instance, in cluster 1, the customers had the highest longevity amongst all, with an average of 682 days (nearly 2 years). These customers also had high F and M levels, while having lower R and P values. These customers can be considered as “loyal and valuable customers”. For the customers in the second cluster, although their L values are high and they are relatively old-timers, but their high R values indicate that they have not been using any service for a long time. So, it can be

**Table 3.** Definition of Occupations.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | internet | 12 | SoundandImage | 23 | Food | 34 | Discussion | 45 | Playvideo |
| 2 | digital | 13 | Books | 24 | Vehicle | 35 | Social | 46 | FlashGames |
| 3 | advertising | 14 | WebDesign | 25 | Children | 36 | Charity | 47 | Recreation |
| 4 | Hosting | 15 | Graphic | 26 | Pets | 37 | Classifieds | 48 | Information |
| 5 | Software | 16 | AdsNetworks | 27 | Medical | 38 | Religious | 49 | Emigration |
| 6 | FileSharing | 17 | Promotional | 28 | Gifts | 39 | Legal | 50 | TravelAgents |
| 7 | SEO | 18 | Deals | 29 | LearningTools | 40 | Securities | 51 | HotelsandTickets |
| 8 | Network | 19 | Hardware | 30 | Courses | 41 | Financial | 52 | TourismServices |
| 9 | Email | 20 | Health | 31 | Marketing | 42 | Estate | 53 | SMS |
| 10 | Charging | 21 | Clothing | 32 | Translation | 43 | Employment | | |
| 11 | Mobile | 22 | HomeandOffice | 33 | Distance Education | 44 | OnlineGames | | |

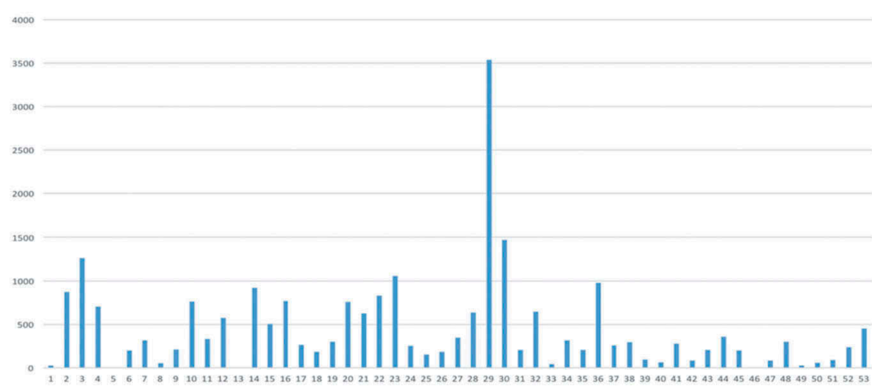


**Figure 2.** Distribution of Occupations for dataset.

**Table 4.** Davis Bouldin Index for the dataset.

| #Clusters | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| Davis Bouldin | 1.32 | 1.07 | 0.87 | 0.92 | 1.09 | 0.95 | 1.00 |
| #Clusters | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| Davis Bouldin | 0.99 | 0.95 | 0.96 | 0.95 | 0.96 | 0.97 | 1.02 |
| #Clusters | 16 | 17 | 18 | 19 | 20 | Optimal Number of Clusters: 4 | |
| Davis Bouldin | 0.93 | 0.99 | 0.98 | 0.91 | 0.93 | | |

assumed that these customers are somehow old and churned. The customers in cluster 3 are those who had a brief relationship with the company and their high R levels indicate that they have lost their interest in using the company's services as well. At last, the customers in cluster 4 can be considered as new and valuable customers given that both their L and R values are low, while they have a relatively high volume of transactions. Moreover, they use the company's services on a regular basis based on P.

### *Cluster 1*

The results illustrate that some groups of customers need more thorough consideration and attention. The mean value of R indicates that some of the customers in this group may also be churned or willing to churn. Also, the mean value of P can be considered as another warning sign. The mean values of F and M indicate that there are very good customers in this group, whose retaining is crucial for the survival of the company.

The number of customers in this group is 5042; a further investigation shows that more than 5,000 of them have been the company's customer for more than one year and 10% of them have been interacting with the organization for nearly 3 years. Approximately 3500 of these customers have used the company's services in the last

**Table 5.** Mean Cluster Information and its appellation.

| | Size | L | R | F | M | P | |
|---|---|---|---|---|---|---|---|
| 1 | 5042 | 682.16 | 36.68 | 3244.19 | 110130554 | 21.83 | Loyal and valuable customers |
| 2 | 1395 | 489.66 | 134.6 | 24.57 | 763942.9 | 131.82 | Old and churned customers |
| 3 | 6391 | 114.78 | *258.7* | 183.6 | 7232646 | 13.95 | Young and churned customers |
| 4 | 10696 | 106.40 | 47.87 | 303.64 | 12319862 | 12.70 | Young valuable customers |

month and nearly 50% in the last 10 days, which indicates that the activity and engagement of this group is pretty well. About 25% of the cluster's population had more than 1000 transactions and over 10% had transactions with the accumulated value of more than 100 million Rials, which have generated a significant liquidity. More than 3,000 people in this group have P less than 20, which indicates that they have been using the services on a regular basis but with slightly overdue periods.

### Cluster 2

These Customers had a long relationship with the company, but the mean value of R shows that they may have churned. It is noteworthy that there were also some among them who had high amounts of transactions and could have been valuable customers. The fact that mean P = 131.82 is also another negative point for this group, which shows that on average, they had a transaction per approximately 132 days.

The number of customers in this group is 1395 and a detailed investigation shows that about 60% of these customers had interactions with the company for more than one year. About 800 of these customers had not any transactions in the last three months (approximately 60%), which may indicate high rates of churning or willing to churn for this group. Only four customers had more than 1,000 transactions, and just one customer had transactions with more than 100 million Rials in value. It is also noticeable that there was no one with $P \leq 50$, which is a negative point for this group. As stated before, this group can be called as the old customers who have left or quitted their relation with the company.

### Cluster 3

Based on Table 5, the mean values of L and R indicate that on average, the customers in this cluster had a short relationship with the company and are churned. However, they had a relatively high value of transaction in the relationship period. The business' that can't survive or operate occasionally can be considered as members of this cluster.

The number of customers in this group is 6391 and only about 10% of these customers have a longevity of over 1 year. In the last three months, none of them had any transactions, and it can be concluded that they have churned. Only 75 people had more than 100 million Rials of transactions, which are accounting for nearly 1% of the group's population. By simultaneously considering the P and R factors, it is likely to assume that these customers have churned.

### Cluster 4

According to Table 5, in this cluster, the average values indicate that there are new customers in this group that conduct financial transactions on a regular basis and their usage period is relatively well. The high mean values of M and F indicate that there are some valuable customers in this group. A more detailed analysis of the group can lead to better market strategies.

The number of customers in this group is 10,696 and the statistical review shows that more than 10,000 of these Customers have been using the services for less than 1 year. Nearly 50% of the customers had transactions in the last 30 days, which reflects the favorable dynamism of this group.

Due to the short longevity of these customers, one can expect that the number of transactions in this group is not so high, but it is noteworthy that about 2% of them have transacted more than 100 million Rials during the period, which is very convenient. About 80% of them have $P \leq 20$, which indicates the dynamism and mobility of this group.

In view of the above explanations, we conclude that clusters 1 and 4 have some customers that are currently in an active relation with the organization. For a better understanding of these individuals, we can also divide the customers of each cluster to obtain more homogeneous populations and develop more appropriate strategies for each sub-cluster.

### Re-cluster

In this section, a new clustering is carried out again using the k-means method and data of cluster 1 and cluster 4. According to Table 6, Davies-Bouldin's index for cluster 1 and cluster 4 has the lowest value when divided into 4 and 5 clusters,

**Table 6.** Davis-Bouldin Index for Cluster 1 and Cluster 4.

| #Clusters | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| Cluster 1 Scores | 1.35 | 1.13 | 0.95 | 0.95 | 0.98 | 1.03 | 1.02 | 0.97 |
| Cluster 4 Scores | 1.31 | 1.1 | 1.08 | 0.97 | 1.05 | 1.04 | 1.01 | 1.02 |

respectively. Therefore, we divide these clusters into sub-clusters and reexamine each of these groups.

As seen in Table 7, which includes the mean of variables, all individuals in cluster 1 are old customers. But the cluster 1_4 with an average longevity of more than 3 years has included the most loyal customers of the company. These customers recently had transactions and, on average, they have the highest values of F and M factors. This necessitates making appropriate strategies to retain them because these are loyal and valuable customers.

The cluster 1_3 also includes customers who have recently used the services of the company and have a roughly regular basis with short periods in using the company's services. The amount of money transacted and the frequency of use of the company's services in this group indicates that this group consists of valuable customers as well, whose retention is of great importance. The cluster 1_2 includes customers who have less value for the company compared to the clusters 3 and 4, but their longevity shows that these customers are loyal.

For the customers in the cluster 1_1 the company needs to put in place necessary and efficient strategies; because these old customers are leaving the company's services or have already left. The amount of money transacted and their high number of transactions indicate their worthiness, but regarding the values of P = 25 and R = 159, one can find out that they are churning.

On the other hand, population of cluster 4 is 10696 customers, which makes it the largest cluster in the primary clustering. On the other hand, these customers are often new to the services, which raises the important issue of making the appropriate strategies tailored for them. By doing so, the company can convert a fair amount of them to high-valued customers or, at least, prevent them from churning.

### *Cluster 1_1*

The population in this cluster is 665 customers. The company interaction lifespan for almost all individuals in this cluster is more than 1 year. Only 14 people have used the company's services in the last 90 days. Given that the maximum value of P is 81, it can be concluded that a large number of these people are churned so there is an immediate requirement to provide appropriate strategies to restore them. Furthermore, 7% of the customers in this group had more than 100 million Rials and 31% more than 10 million Rials worth of transactions; These customers did not use the services on a regular short-term basis ($P \leq 5$), but all of them had at least one transaction in 100-day intervals.

### *Cluster 1_2*

These customers are considered as old ones. The minimum of L variable in this cluster is 343, which indicates that almost all members have been using the services for more than one year. The number of members in this cluster is 808. Most of the customers have used the services in the last 3 months, but the average transaction value is not considerable, as well as the number of transactions. These customers can be mentioned as the medium value customers. The main issue is that considering the seniority of these people, whether it is possible to induce some customers from this group to become valuable senior customers.

### *Cluster 1_3*

Highly valued customers are the general feature of this cluster. The number of customers in this cluster

**Table 7.** Mean value of Sub-clusters of cluster 1 and Cluster 4 Information and appellations.

| | Size | L | R | F | M | P | |
|---|---|---|---|---|---|---|---|
| 1_1 | 665 | 673.87 | 159.35 | 2242.69 | 46597609.59 | 25.61 | Old and churned customers |
| 1_2 | 808 | 591.92 | 37.95 | 74.29 | 2951118.29 | 53.63 | Loyal and medium value customers |
| 1_3 | 3056 | 620.35 | 12.78 | 2452.98 | 76940464.97 | 13.58 | Loyal and valuable customers |
| 1_4 | 513 | 1203.27 | 18.31 | 14248.47 | 559018087.4 | 16.05 | Loyal and valuable customers |
| 4_1 | 2642 | 32.90 | 110.77 | 69.37 | 3735876.54 | 5.92 | Young and churned customers |
| 4_2 | 2352 | 262.37 | 18.85 | 681.13 | 27709793.7 | 13.99 | high value customers |
| 4_3 | 4365 | 39.75 | 21.51 | 108.38 | 4544058.55 | 6.71 | Young and medium value customers |
| 4_4 | 26 | 161.15 | 25.96 | 36641.88 | 1353759425 | 7.38 | Young and high value customers |
| 4_5 | 1311 | 195.5 | 61.36 | 27.99 | 1294480.4 | 44.12 | Young medium value customers |

is 3056. They can be referred to as highly valued customers because most of them had transactions in the last 3 months, have a longevity of nearly 1 year, and more than half of them have conducted transactions worth more than 10 million Rials. The company has to provide appropriate strategies for this group as well, to retain them and enhance the customers' value, so that in addition to keeping them loyal, it can make more customers become valuable.

### *Cluster 1_4*

In this group, there are customers who, in addition to having a long longevity, have recently used the company's services and also have a high transaction volume and high transaction frequency. This group can be called the most valuable group in the main cluster (cluster 1) because of their relationship length with the company and high monetary value. There is a vital need to make strategies to retain these customers. Of these 513 people, more than 21% had transactions worth more than 100 million Rials, and almost all of them were customers for over 2 years. Considering the P factor, almost all of these individuals are engaged in financial transactions during 2-month intervals on average, and since more than 89% of the customers have $R \leq 60$, they are not only highly valuable, but also loyal and regular.

### *Cluster 4_1*

There are 2642 customers in this group. These customers can be counted as young but churned customers, for the following reasons: first, the longevity of more than 94% of them has been less than 5 months. Moreover, over 78% of them had not any transactions in the last 3 months. In fact, the customers in this category are mostly those who have not used the services after a few times. The variable F was higher than 50 for only 384 customers, which indicates the low number of transactions. It is noteworthy that, there are 19 customers in this category who had transactions worth more than 100 million Rials in a short period of time, which points to the fact that the company can preserve these customers and help them become high-valued and loyal ones by adopting appropriate strategies for them.

### *Cluster 4_2*

There are 2352 customers in this cluster, which nearly 96% have been using the services for more than 5 months, and over 77% of them had interactions with the company in the last month. Also, 149 of these customers had transactions worth over 100 million Rials in this short period, which puts them among high-valued customers. These customers are purchasing with a period of roughly 20 days, which is another good sign for the company that they can be considered as regular and highly active customers.

### *Cluster 4_3*

With a total of 4365 customers, this group contains the largest share of customers in the main cluster 4. These customers are young and medium valued. Only 22 customers have a longevity of more than 5 months, and 35 individuals had transactions worth more than 100 million Rials. Most of these customers have an average transaction period of 20 days, and their last transaction was less than 60 days old, which indicates that they can be considered as young and active customers.

### *Cluster 4_4*

This group is consisted of only 26 customers. These customers are young and yet very high valued ones. About 50% of them have been using the services for more than 5 months. By studying the two factors of R and P at the same time, it is suggestive that three customers should be monitored more accurately in order to determine whether they are churned or being fraudulent clients. All the members in this group have transactions worth more than 100 million Rials, which is a sign of importance of this group.

### *Cluster 4_5*

This group has a relatively medium longevity and contains 1311 customers. Different parameters for this group indicate that all of their factors are average, which points to the issue of adopting appropriate policies for encouraging them to carry out more transactions.

## Preliminary conclusions

Understanding the business customers' behavior is one of the more important issues and recent agendas that organizations should take into account in their interactions with their business customers. This knowledge can help organizations to ensure the efficiency of their marketing, relationships and customer management. In this practitioner note, by studying the L, R, F, M, and P criteria, the customers of a B2B Fintech industry are segmented and analyzed. The results indicate that the L and P variables can significantly help with the interpretation of and tailoring marketing strategies towards each cluster of customer groups. Furthermore, a hierarchical clustering approach has been used to gain more in-depth understanding and efficiency of marketing initiatives towards extant customers. The preliminary results of this research demonstrate that it is possible and important to find specific segments within more general segments of customers that follow particular patterns and behaviors; therefore, the adoption of different strategies for these segments can lead to more accurate knowledge, understanding, and management of these customer segments' tendencies.

The contributions of this practitioner note are threefold; first, the literature of the RFM approach towards customer segmentation have further developed by this research building on the L and particularly the P factors. The P factor has only recently been adopted in the segmentation realm by the work of Peker, Kocyigit, and Eren (2017) in the grocery retail industry. Second, this practitioner note extends the application of the LRFMP model in the B2B setting in general and the Fintech sector in particular. In addition to the efficiency brought about deploying this approach, it demonstrates, once again, the applicability of customer relationship and segmentation methods in the B2B setting; and third, the two-stage approach adopted in this practitioner note not only adds further precision to the clustering of customers, but also subsequently enhances the efficiency of marketing strategies towards them.

With respect to limitations, it is worth considering that the data is drawn from the customers who were active in the last 12-month period and one product only. Future research can build upon a more comprehensive body of data with more products and times frames extended. Further studies can also take into account the implications of marketing strategies on the behavior of each segment. This adds further dynamism to the process of customer segmentation and understanding. It is also recommended to include more data – e.g. demographic and industrial sectors – of customers in the process altogether. Adding an element of customer referrals can also be considered as yet another factor in the evaluation of customer behavior. Last, but not least, future studies can predict customer behavior building on this process and provide a more holistic framework.

## References


Chen, D., S. L. Sain, and K. Guo. 2012. Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. *Journal of Database Marketing & Customer Strategy Management* 19 (3):197–208. doi:10.1057/dbm.2012.17.

Chen, K., Y. H. Hu, and Y. C. Hsieh. 2015. Predicting customer churn from valuable B2B customers in the logistics industry: A case study. *Information Systems and e-Business Management* 13 (3):475–94. doi:10.1007/s10257-014-0264-1.

Colombo, R., and W. Jiang. 1999. A stochastic RFM model. *Journal of Interactive Marketing* 13 (3):2–12. doi:10.1002/(SICI)1520-6653(199922)13:3<2::AID-DIR1>3.0.CO;2-H.

Davies, D. L., and D. W. Bouldin. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2):224–27. doi:10.1109/TPAMI.1979.4766909.

Dibb, S. 1998. Market segmentation: Strategies for success. *Marketing Intelligence & Planning* 16 (7):394–406. doi:10.1108/02634509810244390.

Hosseini, S. M. S., A. Maleki, and M. R. Gholamian. 2010. Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications* 37 (7):5259–64. doi:10.1016/j.eswa.2009.12.070.

Hsieh, N. C. 2004. An integrated data mining and behavioral scoring model for analyzing bank customers. *Expert Systems with Applications* 27 (4):623–33. doi:10.1016/j.eswa.2004.06.007.

Jonker, J. J., N. Piersma, and R. Potharst. 2006. A decision support system for direct mailing decisions. *Decision Support Systems* 42 (2):915–25. doi:10.1016/j.dss.2005.08.006.

Kandeil, D. A., A. A. Saad, and S. M. Youssef 2014, September. A two-phase clustering analysis for B2B customer segmentation. In *Intelligent Networking and Collaborative Systems (INCoS), 2014 International Conference on* (pp. 221–28). IEEE: Salerno, Italy.

Khajvand, M., K. Zolfaghar, S. Ashoori, and S. Alizadeh. 2011. Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. *Procedia Computer Science* 3:57–63. doi:10.1016/j.procs.2010.12.011.

Kim, J., E. Suh, and H. Hwang. 2003. A model for evaluating the effectiveness of CRM using the balanced scorecard. *Journal of Interactive Marketing* 17 (2):5–19. doi:10.1002/dir.10051.

King, S. F. 2007. Citizens as customers: Exploring the future of CRM in UK local government. *Government Information Quarterly* 24 (1):47–63. doi:10.1016/j.giq.2006.02.012.

Li, D. C., W. L. Dai, and W. T. Tseng. 2011. A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business. *Expert Systems with Applications* 38 (6):7186–91. doi:10.1016/j.eswa.2010.12.041.

Li, S. T., L. Y. Shue, and S. F. Lee. 2008. Business intelligence approach to supporting strategy-making of ISP service management. *Expert Systems with Applications* 35 (3):739–54. doi:10.1016/j.eswa.2007.07.049.

Li, Y. M., C. H. Lin, and C. Y. Lai. 2010. Identifying influential reviewers for word-of-mouth marketing. *Electronic Commerce Research and Applications* 9 (4):294–304. doi:10.1016/j.elerap.2010.02.004.

Ngai, E. W., L. Xiu, and D. C. Chau. 2009. Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications* 36 (2):2592–602. doi:10.1016/j.eswa.2008.02.021.

Peker, S., A. Kocyigit, and P. E. Eren. 2017. LRFMP model for customer segmentation in the grocery retail industry: A case study. *Marketing Intelligence & Planning* 35 (4):544–59. doi:10.1108/MIP-11-2016-0210.

Sohrabi, B., and A. Khanlari 2007. Customer lifetime value (CLV) measurement based on RFM model. 10.1094/PDIS-91-4-0467B

Spring, P., P. S. Leeflang, and T. Wansbeek. 1999. The combination strategy to optimal target selection and offer segmentation in direct mail. *Journal of Market-Focused Management* 4 (3):187–203. doi:10.1023/A:1009899802421.

Wu, H. H., E. C. Chang, and C. F. Lo. 2009. Applying RFM model and K-means method in customer value analysis of an outfitter. In: Chou SY., Trappey A., Pokojski J., Smith S. (eds) *Global Perspective for Competitive Enterprise, Economy and Ecology*. Advanced Concurrent Engineering. London: Springer.