Customer Value Analysis Using Weighted RFM model: Empirical Case Study

Tarek BELHADJ*

Abdelhafid Boussouf University Center,
Mila, Algeria

t.belhadj@centre-univ-mila.dz

Received: 11/03/2021 *Accepted*: 12/10/2021 *Published*: 16/12/2021

Abstract :
The aim of this article is to provide a comprehensive methodology to perform customer value analysis. First, we review the CLV concept, RFM model and other related concepts. Then a case study for transactions database of a chosen company is presented. We use Weighted RFM (recency, frequency and monetary) model to determine customers' lifetime value (CLV) and segment them based on k means clustering approach. The relative weights of RFM model were determined by analytic hierarchy process (AHP). Findings revealed that CLV analysis using RFM measures can help the company better understand its customers and locate the most profitable customers. Therefore, the company can apply marketing strategies to maintain customer relationships more effectively.
***Key Words*:** Analytical Hierarchy Process; Customer Lifetime Value; Customer Relationship Management; K-Means Clustering; RFM model.

JEL Classification : M31, C38, C69.

* Corresponding author: Tarek Belhadj (*t.belhadj@centre-univ-mila.dz*)

Introduction:

Many studies have proved that attracting new customers is more costly than retaining old ones (Larivière & Van den Poel, 2005; Hwang, Jung, & Suh, 2004). It has been estimated that obtaining new customers costs five times more than retaining an existing one (Rust & Zahorik, 1993). From the perspective of customer relationship management, customers are not equal in terms of their values to the company. Therefore, evaluating customers values in order to understand their profitability and retain valuable customers became a critical parts of customer relationship management activities (Shen & Chuang, 2009). Companies are increasingly recognizing the importance of customer value analysis to identify profitable customers and to develop strategies to target customers. The progress of information technology and the growing availability of customer transaction data enabled companies to have an important base for decision making. Such a huge volume of data (Big Data) is useless unless implementing data analytics plans which enable companies connecting customers to their marketing decisions (Verhoef, Kooge, & Walk, 2016). The literature of customer relationship management suggests diverse and rich approaches of customer value analysis.

Among it, Customer lifetime Value (CLV) approach based on Recency, Frequency, Monetary (RFM) model represents an important alternative, because it requires a very small number of variables, it is easy to understand and implement and it is very effective in identifying valuable customers which can help the company boosting its profit in a short time (Wei, Lin, & Wu, 2010). Recently, some researchers proposed an enhanced Weighted RFM model. They devoted different weights to each of RFM variables depending on business features and characteristics of the industry (Shih & Liu, 2003; Shen & Chuang, 2009; Wei-Jiang, Shu-Yong, Xue, & Xiao-Feng, 2011). The aim of this clustering and analytic method is to identify the most profitable customers and enable decision makers to allocate resources and make marketing strategy more effective.

Research questions: In view of the above discussion, the research questions are as follows:

(i) What is customer value analysis and how is it applied?

(ii) Can RFM analysis be used to segment and identify the most profitable customers?

(iii) What is customer CLV and how is it estimated using RFM value?

Research methodology: This article aims at implementing a CLV analysis based on RFM model. For this purpose, the framework of the paper will be as follows. First, we review the CLV concept, RFM model and other related concepts. Next, the application of these concepts is presented using a case study for transactions database of a C company. Then, the analysis outputs and result discussion are presented. Finally, the paper ends with concluding remarks.

I. Literature Review:

This section mainly explores the study background. It reviews related studies of customer value analysis, RFM model, classifying algorithm and Customer Lifetime Value (CLV).

Customer value is one of the most important key concepts in customer relationship management CRM. Today's companies are very interested to know the value of their deferent customers. Customer value analysis is "a kind of analytic method for discovering customers' characteristics and makes a further analysis of specific customers to abstract useful knowledge from large data" (Cheng & Chen, 2009, p. 4177). From the CRM perspective, customers are not homogeneous. Therefore, market segmentation is a necessary tool to deal with customer diversity. It is a process of dividing customers into distinct and homogeneous groups in order to develop differentiated marketing strategies based on their characteristics (Kadir & Achyar, 2019). For a long time, customer segmentation models were based on traditional criteria such as: demographic, geographic, and psychographic features of customers. Despite its importance, traditional segmentation fails to consider major shifts in today's complicated business environment caused by the advance of Information Technology. Thus, customer data collected by sophisticated information systems and advanced analytics technique support more accurate customer segmentation based on transactional and behavioural data (Lee & Park, 2005; Huseynov & Yıldırım, 2017). Methods such as Customer Lifetime Value (CLV) analysis; Recency, Frequency and Monetary (RFM) model and Customer Pyramid have been developed as important marketing tools that help companies analyse the profitability of its customers and improve the customer segmentation to customize its marketing strategies.

Over the past twenty years, several studies have incorporated RFM model and Customer Lifetime Value (CLV) to improve customer value analysis and develop accurate prediction and classification models. For example, Shih & Liu (2003) presented a systematic approach in evaluating customer lifetime value (CLV) by means of analytic hierarchy process (AHP) to determine the relative weights of RFM variables. The proposed approach was applied to marketing database of a hardware retailer. Clustering analysis was employed to segment customers based on weighted RFM value. The study also discussed three viewpoints for validating the proposed method.

Cheng & Chen (2009) with the aim to improve segmentation accuracy and enhance customer classification rules using data-mining model, they suggested a new procedure that includes quantitative value of RFM attributes and K-means algorithm integrated with rough sets theory. Empirical case study was performed to validate the proposed procedure. The finding revealed that the proposed procedure is more efficient than the listed methods in terms of accuracy rate in classifying the segmentation of customer value and the output of proposed procedure represent understandable decision rules.

Kumar, Chaitanya, & Madhavan (2012) focused on clustering e-banking customers using RFM approach model. The aim was to improve the relation between marketing decision and customer segmentation. The study analyses customer characteristics and behaviours with appropriated criteria: access time, transaction access and RFM Analysis, Life Time Value of the customers (LTV), demographic variables. The analysis procedures consisted of two phases. Firstly, customers were segmented into clusters according to their RFM values using K-Means clustering. Secondly, the resulted clusters were again divided into new clusters based on demographic data. Finally, LTV analysis was used to identify customer's profile.

He & Li (2016)proposed a new customer segmentation approach based on three dimensions namely customer lifetime value, customer satisfaction and customer activity. The appropriated variables were obtained using RFM model, Kano model and BG/NBD model. The study concluded that, decision-makers can use the output of customer segmentation to better identify market segments and developed more effective marketing strategies.

Christy, Umamakeswari, Priyatharsini, & Neyaa (2018) have performed a segmentation process on a transactional dataset of the customers of a company based on RFM model. Then the clustering analysis was extended using K –Means algorithms, Fuzzy C – Means and a new proposed algorithm RM K-Means. The different approaches were compared with one another. The results revealed that the new proposed algorithm is more effective in terms of its iterations and execution time.

1. Customer Value Analysis:

As we have mentioned above**,** customer value analysis is an analytic method for abstracting customers' characteristics by using transaction database and then enhance the customer relationship management. It is clear that customers are not homogeneous in their purchasing behaviours. Therefore, they vary extensively in the value they represent to the company (Cheng & Chen, 2009). Thus, companies aim via using value analysis method to know their customers who spend more money and make the most contribution in the profits. Since, the literature suggest that the cost of customer retention is far less costly than acquisition cost, customer value analysis is used as an effective tool to identify profitable customers thus helping decision-makers to decide which customers to give particular marketing strategies.

There are several methods which are used in customer value analysis. Pareto analysis and the Pyramid model are discussed below. RFM and CLV will be discussed further later.

1.1. Pareto analysis:

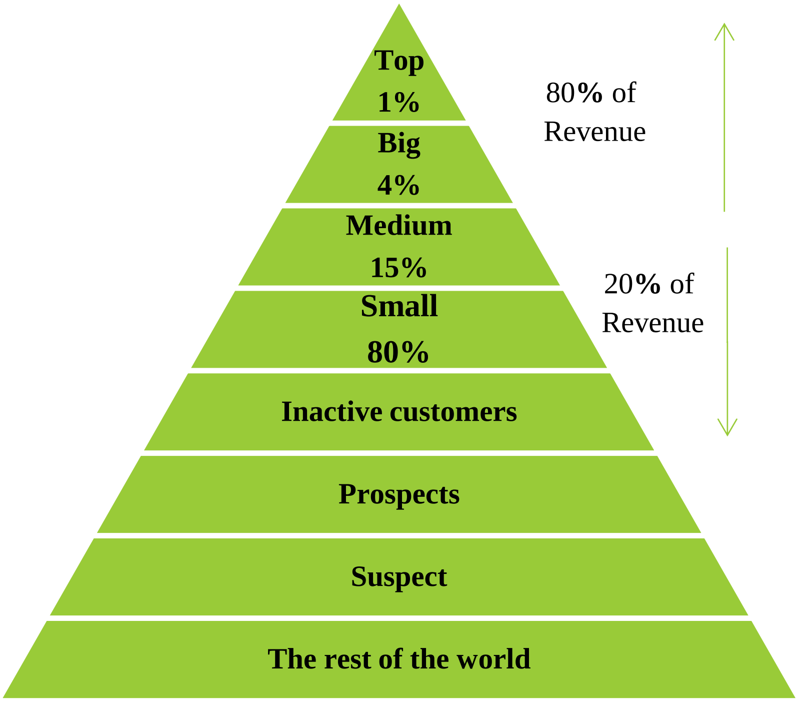
The Pareto Law, named after the well-known Italian economist and sociologist Vilfredo Pareto (1848-1923), states that for many events, roughly 80% of consequences come from 20% of the causes. Marketing literature follow Pareto's 80/20-law suggested that 20 % of customers generate 80% of revenue (Ultsch, 2002). This can be used to prioritize customers which result optimal output. Therefore, particular marketing strategies can be elaborated to retain the top 20% customers and boost the lowest 80% customers towards the top 20% .

1.2. Customer Pyramid :

One of the most used analysis methods of customer profitability is the Customer pyramid. Curry & Curry (2002) introduce the customer pyramid as way to segment clients according to revenue. The active customers are divided into four segments: Top, Big, Medium, and Small as shown in Figure 1. The advantage of Customer pyramid in comparison with traditional market segmentation comes from that the pyramid model help to understand customers' value and identify profitable customers, therefor, attracting and retaining a company's most valuable customers. The pyramid model has been used and proven extremely useful to business companies (Aggelis & Christodoulakis, 2005).

Curry & Curry (2002) suggested in their work that successful application of the pyramid comes to those who follow three-step Customer Marketing Strategy: (1) acquiring new customers into the pyramid; (2) boosting customers higher into the pyramid; (3) keeping the customers in the pyramid.

Fig. 1: The Customer pyramid



Source: Curry & Curry (2002, p. 9)

2. RFM Model:

2.1. RFM model definition:

The RFM model is proposed by Hughes (1994), and it is one of the most popular customer value analysis methods. Its advantage is to extract characteristics of customers by using fewer criterions (three variables – Recency, Frequency, and Monetary) as cluster attributes so that reduce the complexity of model of customer value analysis.

The definitions of RFM criterions are described as follows:

Recency (R) refers to the interval (number of days) between the last purchase and the present time. Recency is often regarded as the most important measure from the three criterions, because the most recent purchasers are likely to purchase again, so they potentially create future value. Frequency (F) is the number of transactions / purchases made by the customer in a particular period. The higher the frequency is, the greater customer loyalty is supposed. Monetary (M) represent the amount of money spent by a customer in a specific period of time (Wei, Lin, & Wu, 2010; Mesforoush & Tarokh, 2013)

RFM is considered as one of the most important models used for market segmentation that distinguish important customers and identify customer's purchase behaviour. With the RFM framework, customers can be grouped and classified into segments according to their RFM scores. So, the aim of RFM scoring is to drive better market segmentation that helps companies expect future customer behaviour. Therefore, the quantification of consumer behaviour plays a critical role in interpreting customer value through time.

2.2. Weighted RFM model:

Although the easy scheming of the RFM variables, whether by scores or numeric values, there are two views to calculate a single RFM value. Hughes (1994) and Stone (1995) suggested two opinions which differ, one from the other, in regard to the weight of the three variables of RFM model. . Hughes (1994) considered that the three variables are equal in term of their weights and they have the same importance when calculating a combined score. In the contrast, Stone (1995) stated that the weights of the three variables are not the same. They have different importance depending on the product features and industry characteristics. Thus, different weights must be given to each element of RFM for computing a single RFM value. For instance, Stone (1995) proposed ordering the three variables as follows: Frequency with highest weight, followed by the Recency, then Monetary value with the lowest weighting. However, Khajvand, Zolfaghar, Ashoori, & Alizadeh (2011) and Monalisa, Nadya, & Novita (2019) obtained an RFM weights values that put frequency at the higher importance, followed by monetary and then recency. Other studies found Monetary to have the most value and Recency to have the least value (Shen & Chuang, 2009; Mesforoush & Tarokh, 2013).

Moreover, Miglautsch (2000) suggested a formula of calculating the total RFM score as follows = (R×3) + (F×2) + (M×1). In contrast to the arbitrarily formulas of Miglautsch (2000), Shih & Liu (2003) applied an approach to determine the different weights of RFM variables using Analytic Hierarchy Process (AHP). AHP is a "measurement theory through pairwise comparisons and depends on expert judgment to get priority scales" (Saaty, 2008, p. 83).

2.3. Analytical hierarchy process:

In this study, hierarchical analysis process, one of useful tools in method selection, is used for weighing the relative importance of RFM variables. AHP was developed by Thomas L. Saaty in the 1970s and has been since then extensively used for weighted scoring decision making in complicated problems where many variables or criteria are considered in priority setting. Mainly, the AHP is based on the subjective experience and knowledge of the decision makers to get priorities of variables. Decision makers score the different variables using the comparison of pairs to determine the best solution. This comparison can be done using the scale of relative importance suggested by Saaty (2008). To determine the relative importance of a variable, decision makers or users attribute values that vary from 1 to 9, when compared to another variable, as seen in Table 1.

Table 1: Scores for the importance of variable

Importance Scale	Definition of Importance Scale
1	Equally Important Preferred
2	Equally to Moderately Important Preferred
3	Moderately Important Preferred
4	Moderately to Strongly Important Preferred
5	Strongly Important Preferred
6	Strongly to Very Strongly Important Preferred
7	Very Strongly Important Preferred
8	Very Strongly to Extremely Important Preferred
9	Extremely Important Preferred

Source: Taherdoost (2017, p. 245)

The AHP is performed in three main steps (Shih & Liu, 2003; Shen & Chuang, 2009) as follow:

(1) Perform pairwise comparisons: evaluators (decision makers) were asked to do paired comparison of the relative importance of RFM variables, and give the value of 1 to 9 to each indicator using the scale as shown in Table 2.

Table 2: Sample AHP Questionnaire

Idicator	◄								►	**Idicator**
Rcency	**9 .**	**7 .**	**5 .**	**3 .**	**1 .**	**3 .**	**5 .**	**7 .**	**9**	Frequency
Monetray	**9 .**	**7 .**	**5 .**	**3 .**	**1 .**	**3 .**	**5 .**	**7 .**	**9**	Frequency
Monetray	**9 .**	**7 .**	**5 .**	**3 .**	**1 .**	**3 .**	**5 .**	**7 .**	**9**	Rcency

(2) Assess the consistency of pairwise judgments: expert evaluators may make inconsistent judgments when making pairwise comparisons. Before the weights are computed based on the pairwise judgments, the degree of inconsistency is measured by the inconsistency index that should be of less than less than 0.1. Otherwise the pairwise judgments may be revised before the weights of RFM are computed.

(3) Computing the relative weights: this step determines the weight of each decision element according to the pairwise comparisons.

3. Cluster analysis:

Clustering is a statistical method that was used for for classifying physical or abstract items into homogeneous segments where the items have similar characteristics. A cluster is a group of data items that are homogeneous and related to the items within this the cluster and , at the same time, are heterogeneous and unrelated to the items in other clusters (Mesforoush & Tarokh, 2013). Cluster analysis is largely used for market segmentation and has been applied in meany studies along with RFM model.

The literature suggested several methods which are used in clustering such as TwoStep algorithm and K-means algorithm. The latter is one of the well-known algorithms for cluster analysis. The K-means algorithm has two major features. Firstly, it is based on the mean value of the objects in the cluster, i.e., the algorithm assigns each item to the cluster with the nearest centroid (mean). Secondly, K-means clustering technique requires predetermination of the number of clusters by the decision maker or the user (Cheng & Chen, 2009; Mesforoush & Tarokh, 2013; Kadir & Achyar, 2019).

1.4. Customer Lifetime Value:

CLV is developed from the literature of customer relationship management (CRM). The aim of CRM is to build and maintain strong relationships with customers and to generate higher customer lifetime value to the company (Estrella-Ramón, Sánchez, Swinnen, & VanHoof, 2013) . CLV can be defined as "the present value of all future profits obtained from a customer over his or hers relationship with a firm" (Gupta, et al., 2006, p. 141). It is essentially used as a basic method to compute and express customer profitability to develop marketing strategies to target the most profitable customers

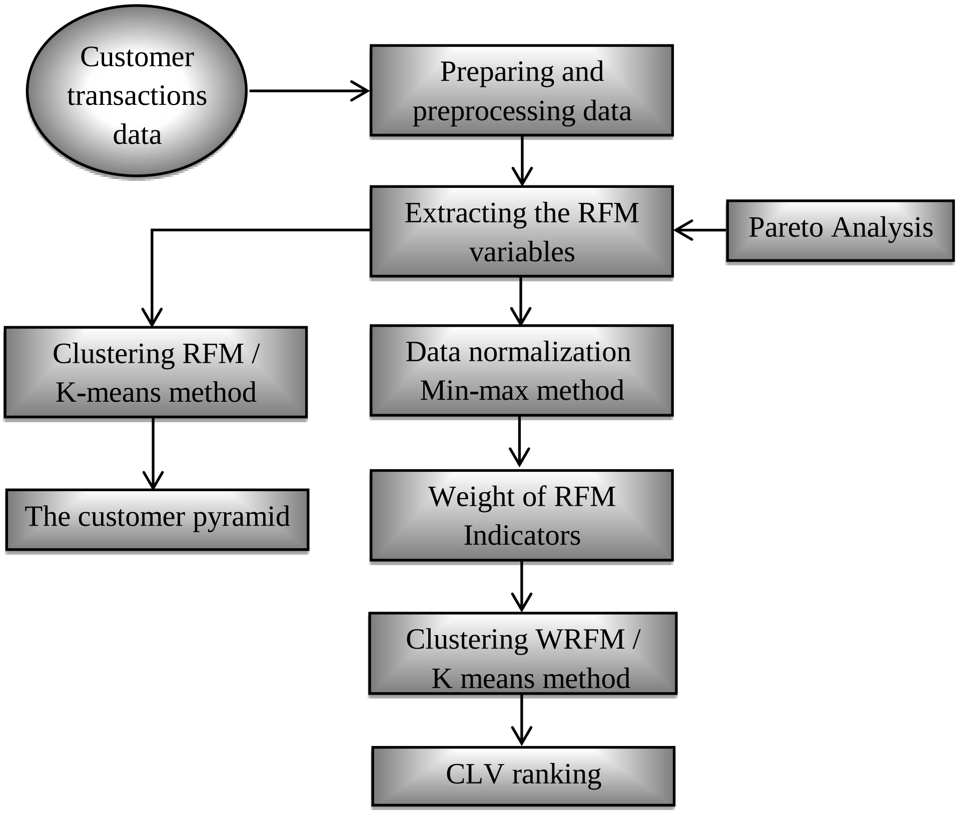
Gupta, et al. (2006) disscussed six implementable modeling approaches that are useful for CLV estimating: RFM Models, Probability Models, Econometric Models, Persistence Models, Computer Science Models and Diffusion/Growth Models. The most powerful and simplest model to implement CLV may be the RFM model – Recency, Frequency, and Monetary value. Gupta, et al. (2006) states that, despite the limitations of RFM-analysis the models remain a mainstay of the industry because of their ease of implementation in practice. Fader, Hardie, & Lee (2005) showed how the well-known RFM (recency, frequency, and monetary value) model can be used with customer lifetime value (CLV) to build a model that overcomes many of its limitations. Therefore, Several studies use RFM model to estimate Customer lifetime value (Miglautsch, 2000; Shih & Liu, 2003; Sohrabi & Khanlari, 2007).

One of the most significant benefit of CLV based RFM model is that the model inputs are nothing more than each customer’s RFM values derived from the same using transaction data. Moreover, there is no need to split the data into two (or more) time periods, the model use the entire customer base as a single sample to estimate the CLV (Gupta, et al., 2006). Calculating the CLV for all of a firm’s customers allows them to categorize customers based on their individual contribution to the organizations profits This helps to develop strategies to deal with each customer differently, instead of treating every customer the same way using the same marketing approaches. Although firms are interested in knowing the current and predicted customer life time value of their customers, they also need to identify the factors they can control that can potentially increase the value of customer. It is not enough to know who are the most profitable customers, it is even more important to determine how to convert currently less profitable customers into more profitable ones (Savoie, 2014).

II. Study Design:

The case study of this article is about (C) company, an Algerian spare parts wholesaler that provides automotive parts and accessories for Asiatic brands. The company is based in Algiers and has two branches, one in Aïn M'lila and the other in Touggourt. This company implemented a registration system of transaction information of customers. The database of consumer transactions contains information about name and customer ID, date of purchase, total price. For the purpose of this study, six months of data on consumer transactions have been selected (from 2019/07/01 to 2019/12/31). It represents 218 customers and 1542 purchase records. RFM values of the 218 customers are extracted from this particular marketing database to measure the customers' CLV following the proposed framework shown in figure 2:

Fig. 2: The research framework



The research steps can be described as follow:

- **Data Pre-processing:** First of all, we collect purchase information from records to get accurate data for analysing. There are 218 raw data of customer transactions of the given case. In each Raw, there is a unique customer with her/his ID, number of transactions, latest transaction date and total amount of money spent in all of her/his purchases at the end of the study period. Therefore, each customer is assigned with three values for recency, frequency, and monetary variables.

- **Pareto Analysis**: in this step, we divide customers into 5 equal parts (quantiles). Then, we test the 20/80 principle (20% of customers contribute to 80% of total revenue
- **Segment customers by RFM model**: The simple sum of RFM values are used to clustering customers using K-Means algorithm. The resulting clusters represent customer segments to be compared with the customer pyramid introduced by Curry & Curry (2002). Furthermore, these customer segments will be analysed in order to identify their different patterns. The RFM values of each cluster is examined whether above or below the overall average value. Based on that, marketing strategy according to cluster pattern can be suggested to have more targeted marketing action.
- **Data normalization**: In order to increase the cohesion of entry types and simplify the analysis the RFM data must be normalised. Min-max normalization method is used. It performs a linear transformation on the original data and makes it lie between 0 and 1 thus bringing all the values RFM model in the dataset to a common scale.
- **Calculating The Weight of RFM Indicators**: this step employs the AHP to evaluate the relative weight (importance) of each RFM variable, and specifically asks decision makers to make intuitive judgments about ranking order to make pairwise comparisons.
- **Weighted RFM clustering:** K-means clustering is then employed to group customers according to the weighted RFM values.
- **CLV ranking:** Finally, the simple weighted sum of normalised RFM values are used to derive CLV ranking and thus customer segments can be identified and compared clearly.

III. Result Analysis:

1. Pareto Analysis:

By conducting Pareto analysis, customers are arranged in descending order based on monetary value (from the best to the worst). The number of customers is then divided into equal quintiles (20%). The results show that 74,48% of sales volume is generated by 20% of the company customers and 80% of customers only contribute to 25,52% of sales. This is a natural trend which is consistent with “Pareto Principle,” states that approximately 80% of all effects come from roughly 20% of the causes. It is obvious that “Pareto’s principle” represents a basic rule in classifying customers based on their profitability. However, the company needs to more identifying its main customers in order to make marketing strategy more effective and achieve the result with minimum effort and cost.

2. RFM evaluation:

In this respect, RFM is considered as one of the most important models used for customer classification that help to identify important customers and their purchase behaviour by three indicators Recency (R), Frequency (F), and Monetary (M). Recency is the interval between the customer's last transaction and the present time reference. Since the R value –by this definition- negatively impacts the RFM score, a reversed form is used to get more consistence. The new R is defined as the number of days between the first date concerned (1/07/2019) and the date of the last customer purchase. For example a date of last purchase on 20/09/2019 is represented by R=82, while a date on 10/12/2019 is represented by R=163. Frequency (F) refers to the number of transactions in a particular period of time. Monetary (M) is the total amount of money spent by the customer over the same period of time. To calculate the RFM Score we have used the formula:

$$RFM\ Score = R+F+M.$$

Table 3 illustrates a part of data input on which RFM Score is calculated:

Table 3: Sample data of RFM for each customer

Customer No	Recency (day)	Frequency	Monetary (DZD)	RFM Score
00001	135	14	1 076 830,00	1 076 979,00
00002	174	7	45 580,00	45 761,00
00003	167	17	101 740,00	101 924,00
00004	46	2	111 500,00	111 548,00
………	….	….	…..….….…	………….
………	….	….	…..….….…	………….
00218	62	7	60 450,00	60 519,00

The next step is to classify customers using the K-means method. This method requires pre-specifying the number of clusters (k). Since we follow the pyramid model, the parameter k is set to 4. We use the K-means clustering algorithm of SPSS to cluster the RFM values. The result of clustering analysis is represented in table 4.

Table 4 shows four clusters with their corresponding number and percentage of customers and the average of recency, frequency and Monetary values within each cluster.

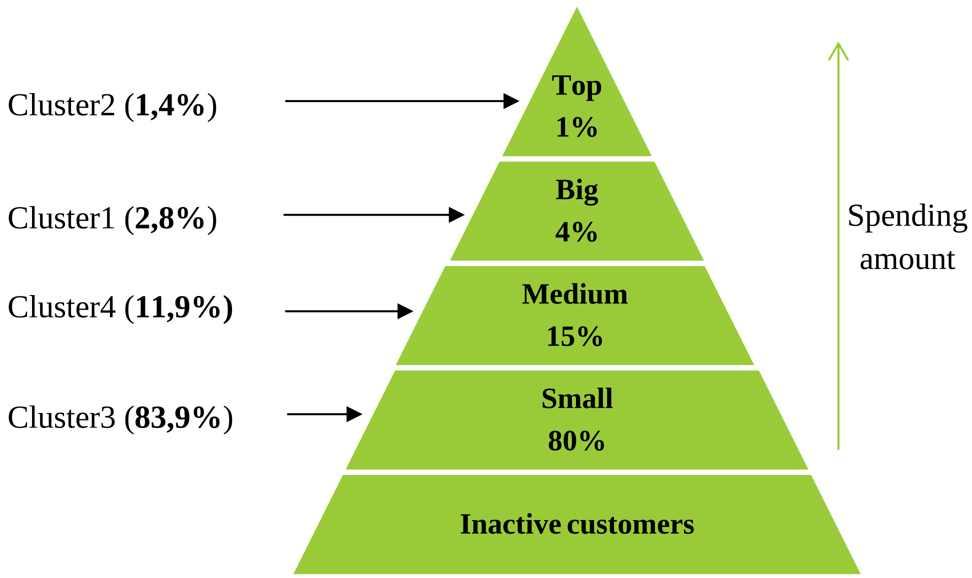
Table 4: The results of RFM K-means clustering for 218 customer

Clusters	Nu. of Customers	Percent	Recency (days)	Frequency	Monetary (DZD)	Pattern
1	6	2,8	135,00	11	1221726,67	R↑ F↑ M↑
2	3	1,4	150,67	11	3878320,00	R↑ F↑ M↑
3	183	83,9	126,89	7	82389,61	R↓ F↓ M↓
4	26	11,9	131,85	7	568996,54	R↑ F↓ M↑
Total / Average	**218**	**100**	**128**	**7,07**	**224020,78**	

3. The customer pyramid:

Following the work of Curry & Curry (2002), we have made a comparison between the above clustering results in table 2 and Curry's Customer Pyramid. The comparison result is shown below:

Fig. 3: Clustering results and Customer Pyramid comparison



As seen in Figure 3, the similarity of the four clusters as derived from the K-means algorithm to the pyramid model is apparent. From the total of 218 customers 183 (83.9% compared to 80%) are included in small segment, 26 (11,9% compared to 15%) in medium segment, 6 (2,8% compared to 4%) in big segment and 3 (1,4% compared to 1%) in top segment.

Furthermore, the resulting clusters represent customer segments to be analysed in order to identify their different patterns. The last line in table 2 also shows the total number of customers and the overall average RFM values for all customers. The last column of the table also shows the RFM pattern of the four clusters. The RFM pattern presents a comparison between the average RFM values of each cluster and the total averages. The comparisons results can be obtained by assigning upward or downward arrow, depending on whether the average RFM values of a cluster is less or higher than the total average.

The pattern of cluster 3 is characterised by low RFM values (R↓ F↓ M↓). This means it has average recency value (the interval between the customer's last transaction and the first date concerned) below the overall average, beside frequency and monetary less than the overall average. This cluster includes customers who had earlier visited the company and made very few transactions with low monetary value. These customers generally make occasional buying and switch companies. To attract such customers, the company should reduce prices and provide extra services. As a result, the company reduces margins and achieves less profit from these customers. So, the cluster 3 includes valueless customers.

Cluster 4 with the pattern (R↑ F↓ M↑) has average frequency below the overall average along with recency and monetary higher than the overall average. This cluster may include new customers who have recently dealt with the company and make important monetary amount. The amount of spending shows that they are tending to gradually increase their purchase frequency. Therefore, the company have to strengthen its relationships with this cluster of customers. They should have special pull marketing strategy to create an ongoing relationship with the company.

Finally, customers in clusters 1 and 2 represent the same pattern. They are characterised by high RFM values (R↑ F↑ M↑) that exceeds the overall average. They represent loyal customers who have a long-term relationship with the company. They have a high purchase frequency and an important contribution to company profitability. Clusters 1 and 2 include high valuable customers. Since, the marketing practices indicates that retaining an existing customer is easier, cheaper and more valuable than attracting a new one, the company should develop marketing strategies to increase customer loyalty in this clusters.

4. Estimating CLV for clusters: To calculate CLV for each cluster, weighted RFM method is used according to the assessments obtained by the AHP. Before that, the RFM scores retrieved from the original database are normalized by the min-max normalization before estimating CLV for clusters.

- **Data normalization:** In the RFM model, R, F, and M have different units. In order to facilitate the analysis the RFM data must be normalized. Min-max method is used to map indicators to the interval [0, 1]. The normalization formula is (Sohrabi & Khanlari, 2007) as follows:

$$R' = \frac{R - Rmin}{Rmax - Rmin} \quad , \quad F' = \frac{F - Fmin}{Fmax - Fmin} \quad , \quad M' = \frac{M - Mmin}{Mmax - Mmin}$$

Where R', F', and M' represent recency, frequency, and Monetary values after normalization. R, F, and M represent the original recency, frequency, and monetary values. Rmax, Fmax, and Mmax are the maximum values of R, F and M for the all customers. Rmin, Fmin, and Mmin are the minimum values of R, F and M for the all customers. All the normalized values are between 0 and 1 and the distribution maintains the same. This process was used also in other studies, where normalized RFM values of each customer were then multiplied by the relative importance of RFM variable, wR, wF and wM (Shih & Liu, 2003; Shen & Chuang, 2009).

- **Weight of Indicators:** To assess the relative importance of RFM variables, hierarchical analysis process is used. So, a company's manager and an academic expert were asked to do paired comparison, and give the value of 1 to 9 to each indicator (see table 2). Using AHP software, the assessments obtained for the relative weights of the RFM variables are mentioned in table 5. The inconsistency value of 0,086 is less than 0.1. The results are reliable.

Table 5: Relative Weight of RFM

Variable	R	F	M
Weight values	0,128	0,087	0,785

- **CLV ranking:** To rank the CLV values we proceed the following steps. The RFM values of each customer are normalised, as described above. The K-means method is then applied to cluster the customers into four groups, according to the weighted RFM values. Then, the CLV value of each cluster is calculated based on the following formula:

$$CLV = wR * R' + wF * F' + wM * M'$$

So, the normalised RFM values of each cluster are multiplied by the relative importance of RFM variable, wR, wF and wM, which are determined by the AHP. The value obtained for each cluster is shown in Table 6.

After that, the CLV ranking of the clusters is determined according to the CLV value of each cluster.

The results in table 6 show that cluster 1 has the highest ranking, it includes the most valuable customers; followed by cluster 3; then cluster 2 in third rank; whereas cluster 4 has the lowest ranking, this later may include the most valueless customers. According to CLV ranking of each cluster, managers of the company have to develop different strategies to attract and maintain customers.

Table 6: The results of RFM K-means clustering for 218 customer

Clusters	Recency (days)	Frequency	Monetary (DZD)	CLV	CLV ranking
1	0,831	0,208	0,974	0,889	**1**
2	0,844	0,114	0,049	0,156	**3**
3	0,845	0,768	0,060	0,222	**2**
4	0,381	0,082	0,027	0,077	**4**
Total Average	**0,705**	**0,126**	**0,055**	-	-

The aim of cluster analysis is to get low similarity (maximising variance) between the different segments while getting high similarity (minimising variance) within these segments. Therefore, analysis of variance (ANOVA) is used to examine whether RFM variables significantly discriminates between the four cluster. The result show significant p-values ($p<0.01$) indicating that RFM variables can use to distinguish between these four clusters.

Conclusion:

The development of information technology enables companies to collect large and important information about their customers. Companies, particularly with increasing sophistication in modelling, can benefit significantly by analysing customer transaction data to determine the most profitable customers and accordingly, adopt the appropriate marketing strategies. Customer lifetime value analysis represents a comprehensive model of customer profitability in customer relationship management (CRM) research. Many methods have been developed to evaluate customer lifetime value.

This article follows the recency, frequency and monetary (RFM) model. Customers are divided into different segments by means of clustering method based on their lifetime value expressed in terms of RFM. The customers data for an Algerian automotive spare parts company has been presented as case study to demonstrate how customer value analysis can be performed using CLV based RFM model. We normalised the extracted RFM values using min-max normalization method and K-Means algorithm was used to cluster customers into four segments (clusters) based on RFM attributes. Then, we employed the analytic hierarchy process (AHP) to evaluate the relative weights (importance) of the RFM variables. Finally, the weighted RFM values were used to calculate customer lifetime values and derive CLV ranking and thus the resulted clusters can be compared clearly and used essentially to explain marketing decisions and CRM strategies by the company.

The paper concluded that a company's adoption of customer value analysis would help better in developing more effective marketing decisions.

Bibliography:

- Aggelis, V., & Christodoulakis, D. (2005). Customer clustering using rfm analysis. In Proceedings of the 9th WSEAS International Conference on Computers (p. 2). 1-5.
- Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. Expert systems with applications, 36(3), 4176-4184.
- Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2018). RFM ranking – An effective approach to customer segmentation. Journal of King Saud University –Computer and Information Sciences. *(In press).*
- Colombo, R., & Jiang, W. (1999). A stochastic RFM model. Journal of Interactive Marketing, 13(3), 2-12.

- Curry, A., & Curry, J. (2002). *The customer marketing method: How to implement and profit from customer relationship management. Simon and Schuster.*
- Estrella-Ramón, A. M., Sánchez, P. M., Swinnen, G., & VanHoof, K. (2013). A marketing view of the customer value: Customer lifetime value and customer equity. South African Journal of Business Management, 44(4), 47-64.
- Fader, P. S., Hardie, B. G., & Lee, K. L. (2005). RFM and CLV: Using iso-value curves for customer base analysis. Journal of marketing research, 42(4), 415-430.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., . . . Sriram, S. (2006). Modeling customer lifetime value. Journal of service research, 9(2), 139-155.
- He, X., & Li, C. (2016). The research and application of customer segmentation on E-commerce Websites. In 2016 6th International Conference on Digital Home (ICDH) (pp. 203-208). IEEE.
- Hughes, A.M. (1994). *Strategic database marketing.*Chicago:Probus Publishing.
- Huseynov, F., & Yıldırım, S. O. (2017). Behavioural segmentation analysis of online consumer audience in Turkey by using real e-commerce transaction data. International Journal of Economics and Business Research, 14(1), 12-28.
- Hwang, H., Jung, T., & Suh, E. (2004). An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry. Expert systems with applications, 26(2), 181-188.
- Kadir, M. A., & Achyar, A. (2019). *Customer Segmentation on Online Retail using RFM Analysis: Big Data Case of Bukku. id.* ICEASD 2019, April 01-02, Indonesia.
- Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. Procedia Computer Science, 3, 57-63.
- Kumar, M. V., Chaitanya, M. V., & Madhavan, M. (2012). Segmenting the banking market strategy by clustering. International Journal of Computer Applications, 45(17), 10-15.
- Larivière, B., & Van den Poel, D. (2005). Predicting customer retention and profitability by using random forests and regression forests techniques. Expert systems with applications, 29(2), 472-484.
- Lee, J. H., & Park, S. C. (2005). x Intelligent profitable customers segmentation system based on business intelligence tools. Expert systems with applications, 29(1), 145-152.
- Mesforoush, A., & Tarokh, M. J. (2013). Customer profitability segmentation for SMEs case study: network equipment company. International Journal of Research in Industrial Engineering, 2(1), 30-44.
- Miglautsch, J. R. (2000). Thoughts on RFM scoring. Journal of Database Marketing & Customer Strategy Management, 8(1), 67-72.

- Monalisa , S., Nadya, P., & Novita, R. (2019). Analysis for customer lifetime value categorization with RFM model. Procedia Computer Science, 161, 834-840.
- Rust, R., & Zahorik, A. (1993). Customer Satisfaction, Customer Retention and Market Share. *Journal of Retailing, 69(2),* 193-215.
- Saaty, T. L. (2008). Decision Making With The Analytic Hierarchy Process. Int. J. Services Sciences 1 (1). 83–98.
- Savoie, P. P. (2014). *A study of customer profitability in automotive retail.Master of Business Administration.* (N. Saint Mary's University (Halifax, Ed.)
- Shen, c. c., & Chuang, H. M. (2009). A study on the applications of data mining techniques to enhance customer lifetime value. WSEAS Transactions on Information Science and Applications, 6(2), 319-328.
- Sheshasaayee, A., & Logeshwari, L. (2017). An efficiency analysis on the TPA clustering methods for intelligent customer segmentation. . *In: 2017 International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, pp. 784–788.*
- Shih, Y. Y., & Liu, C. Y. (2003). A method for customer lifetime value ranking—Combining the analytic hierarchy process and clustering analysis. Journal of Database Marketing & Customer Strategy Management, 11(2), 159-172.
- Sohrabi, B., & Khanlari, A. (2007). Customer lifetime value (CLV) measurement based on RFM model. *Iranian Accounting & Auditing Review,Vol. 14 No. 47, pp 7- 20.*
- Spring, P., Leeflang, P., & Wansbeek, T. (1999). The combination strategy The combination strategy to optimal target selection and offer segmentation in direct mail. Journal of Market-Focused Management, 4(3), 187-203.
- Stone, B. (1995). Successful Direct Marketing Methods, Lincoln-wood. IL: NTC Business Books, 29-35.
- Taherdoost, H. (2017). Decision making using the analytic hierarchy process (AHP); A step by step approach. International Journal of Economics and Management Systems, 2.
- Ultsch, A. (2002). Proof of Pareto’s 80/20 law and Precise Limits for ABC-Analysis. Data Bionics Research Group University of Marburg/Lahn, Germany. 1-11.
- Verhoef, P., Kooge, E., & Walk, N. (2016). *Creating value with big data analytics: Making smarter marketing decisions. Routledge.*
- Wei, J. T., Lin, S. Y., & Wu, H. H. (2010). *A review of the application of RFM model. African Journal of Business Management, 4(19), 4199-4206.*
- Wei-Jiang, Shu-Yong, D., Xue, Y., & Xiao-Feng. (2011). Determination of customer value measurement model RFM index weights. African Journal of Business Management, 5(14), 5567-5572.