K-Means Customers Clustering by their RFMT and Score Satisfaction Analysis

Doae Mensouri, Abdellah Azmani, Monir Azmani
Intelligent Automation Laboratory, FST of Tangier
Abdelmalek Essaadi University, Tetouan, Morocco

***Abstract*—Businesses derive more revenue from building and maintaining long-term relationships with their customers. Therefore, it is essential to build refined strategies based on customer relationship management, with the purpose of increasing their turnover and profits while retaining their customers. In this context, customer segmentation, which is at the heart of marketing strategy, makes it possible to determine the answers to questions relating to the number of investments to be released, the marketing campaigns to be organized, and the development strategy to be implemented. This paper develops an extended RFMT (Recency, Frequency, Monetary, and Interpurchase Time) model, namely the RFMTS model, by introducing a new dimension as satisfaction 'S'. The aim of this model is to analyze online consumer satisfaction over time and discern changes to implement customer segmentation. This article proposes an approach to a segmentation, by client clustering along the unsupervised machine learning method k-means based on data generated using the proposed RFMTS model, in order to improve the customer relationship and develop more effective personalized marketing strategies. The study shows that including satisfaction to the existing RFM model for customer clustering has a major impact and helps identify customers who are satisfied and those who are not, unlike previous attempts to develop new RFM models. By ignoring the "satisfaction" indicator, what went well and what didn't went well cannot be understood. Consequently, the business loses its unsatisfied, loyal, and profitable customers and either fails or relies only on the satisfied ones to continue making profits for an indefinite period of time.**

***Keywords*—Customer segmentation; customer satisfaction; RFMT model; machine learning; k-means**

I. Introduction

As modern economies are primarily service-based, businesses increasingly derive revenue from building and maintaining long-term relationships with their customers. The key to a sustainable e-commerce business is understanding customer characteristics for personalized marketing strategies [1]. In such an environment, marketing aims to maximize customer value [2] and satisfaction, as well as the equity that characterizes the sum of these for all customers of the company.

This study focuses on customer segmentation based on behavioral data due to its availability and evaluation with time. RFM analysis is one of the most renowned techniques used to evaluate customers based on their buying behavior. After collecting and pretreating the data, a new model named RFMTS by adding a new dimension is proposed, which is satisfaction, to the shopping behavior characteristics. Once the values of the new RFMTS model are calculated, k-means the most well-known and widely used clustering method [3], [4] is applied to variables to segment customers. Finally, the behavior of each cluster is analyzed to derive insights and help retailers make the right decisions for each cluster. Grouping customers into different groups on one hand helps to understand customers' needs [5]. On the other hand it allows a company to operate its subsequent segmentation in order to optimize its resources, as well as its sales and marketing efforts [6]. Furthermore, taking customer satisfaction into consideration helps to develop effective marketing strategies and reinforce customers loyalty. Customer satisfaction reflects the difference between the expected and perceived quality of a service [7]. In order to benefit from a large share of the market, companies try their best to reach higher levels of customer satisfaction, which is an important driving force for revenue growth [8], by improving their services [9]. The level of satisfaction determines whether a customer would like to order once again, and become loyal or leave the company [10]. Thus, the motivation for this study to include the satisfaction factor. The application of this approach is made on the e-commerce public dataset, which connects small businesses, but it can be applied to any set of data from which the five RFMTS variables can be extracted.

II. Related Work

A. RFM (Recency, Frequency, Monetary)

RFM was first developed by A. M. Hughes [11] as a method of analyzing customer value. Over the past few decades this models have become a widespread paradigm for behavioral segmentation [12]–[14]. RFM makes it possible to distinguish, at a given moment, customers according to a value which is determined from their time interval since the recording of their last order, their frequency of purchases and the amount spent on their purchases.

- Recency (R): represents the value of the period between a customer's last purchase and current moment. A smaller value of recency implies that the customer purchase frequently. Similarly, a big value implies that the customer won't make a purchase shortly.
- Frequency (F): represents the number of purchases made by the customer. It is equal to the total number of purchases. A high value of frequency implies a high level of customer loyalty.

- Monetary value (M): represents the average value of a given customer's purchases. This is equal to the sum of all customer's purchases divided by the total number of purchases. A customer with a high monetary value, is a customer who provides high revenue to the company.

Several publications have attempted to develop new RFM models, taking into account additional variables, to test whether they perform better than the traditional RFM model. For example, Yeh et al. [14] extended the RFM model, namely RFMTC, by adding two parameters time since first purchase (T) and churn probability (C). In 2012, Alvandi et al. and Wei et al. [15] have developed LRFM model taking into account the duration (L) between the first and the last purchase of a customer. Based on the previous model, Peker et al. proposed a new one (LRFMP) by including the customer visit period (P) [16]. In 2017, Moghaddam et al. developed a new RFM model, namely RFMV, by adding a variable of variety of products (V) [17]. In 2018, Yoseph and Heikkila [18] introduced purchase rate of change (C) to show the quantity and sign of change in customer buying behavior. In 2021, a new parameter was introduced, the inter-purchase Time (T), defined as the time interval between two consecutive purchases by a customer in the same store or on the same website [1]. However, these expanded RFM models lack information revealing customer satisfaction ratings. In this regard, in this paper, RFMT model is extended as RFMTS model by taking satisfaction (S) into account.

B. Customer Satisfaction

The existence of many service providers depends heavily on their competitive advantage in terms of customer satisfaction (CS) [19]. Existing studies [20], [21] have argued that modeling CS is an important research topic. Fornell defined CS as the overall evaluation of the quality of a product or service by a customer based on his purchase and consumption experience [22]. While R. L. Oliver [23] defined it as a judgment of the comfort level of a product/service feature or the product or service itself provides (or is providing) satisfaction related to consumption, including levels of underachievement or overachievement. Customer satisfaction has been examined extensively in past research, which includes the impacts of customer satisfaction [24], [25] and identify the determinants of CS [20], [26]. Customer satisfaction has not been introduced into the extant RFM model for customer segmentation. This article intended to introduce Satisfaction into the RFMT model to create a better performed model, namely RFMTS. Natural language processing is generally used to increase knowledge concerning customer satisfaction [27]. In this study S is equal to the sum of all customer's review score divided by the total number of reviews, since the review score is already defined in the database. Let RS_0 ... RS_n be the review scores given by a customer c, thus the Sc value can be calculated as follows (Equation 1).

$$Sc = \frac{\sum_{i=0}^{n} RSi}{n+1} \quad (1)$$

C. Decision Making

RFMTS model can be used to segment customers in order to identify which customers are satisfied, unsatisfied, active, promising, lost, basic, low-value purchasers, loyal high spenders, at risk. Understanding segments can help retails better tailor their products, marketing strategies and investments. Several combinations of R, F, M, T, S are possible, some canceling out in front of the others. Table I shows the characteristics of the segments and the actions to take into account when dealing with some of them.

TABLE I. SEGMENTS CHARACTERISTICS

Segment characteristics	**Action**
High R, High F, High M, Short T, High S	Loyal customers: improve their value by enticing them to place a few more orders. Increase their loyalty by offering exclusive perks that only the best customers can access.
High R, High F, High M, Short T, Low S	Loyal unsatisfied customers: solve their problems as fast as possible, can't lose them at any cost.
High R, Low F, Low M, Short T, High S	Promising customers: turn them into loyal customers by earning their trust, providing great service, and making them lives easier.
Weak R, Low F, Low M, Long T, Low S	Lost customers: identify their cause of dissatisfaction, learn from it and design campaigns to reactivate them
High R, Low F, High M, Moderate T, High S	Active customers: just like basic ones, keep them interested and active by personalizing their offers.
Weak R, Low F, Moderate M, Short T, Hight S	Basic customers: maximize their value by personalizing the offers intended for them. Take responsibility and offer a sincere apology if an error has occurred

III. IMPLEMENTATION

A. Data Origin

The application of the approach presented in this article, is based on a dataset provided by Olist Store, the largest store in Brazilian markets that connects small businesses. Therefore, this dataset contains many purchases made between the periods 2016-2018 in different stores that used the Olist service. It is an open-source database, downloadable from the Kaggle site (https://www. kaggle.com), known as "Brazilian E-Commerce Public Dataset" and whose Fig. 1 illustrates its model with its different objects and their associations. In the application of the approach presented here, only the Payments, Customers, Orders and Reviews sub datasets (described by Tables II, III, IV and V) are used because they contain the variables necessary for the different variables' calculations.

Initially the number of rows of the Customers, Orders, Payments, and Reviews sub datasets is 99.441, 99.478, 99.478 and 98.410 respectively. A first filter was performed to keep only the orders already concluded (order_status column) reducing their size to 96,478. For the customer data subset, after a first filtering only clients of the concluded orders were kept, its cardinality is no more than 93357. As for the payment, it had to be adapted to that of the orders with the same number of elements.

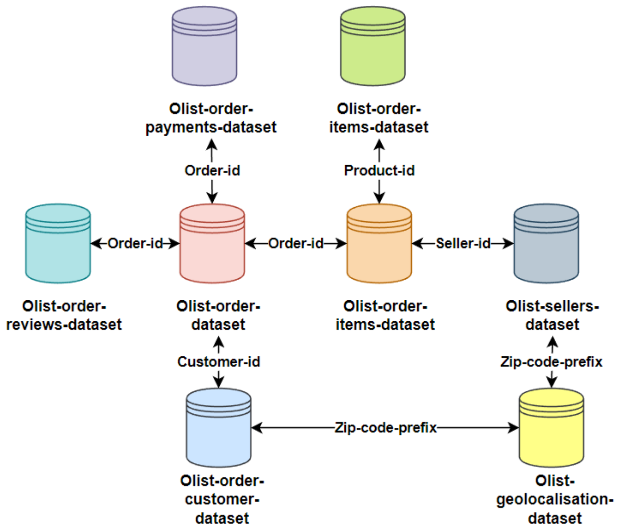

Fig. 1. The Dataset.

TABLE II. PAYMENTS DATASET

Column	Description
Order_id	Order id
Payment_sequential	Sequences of the payments made in case of EMI
Payment_type	Mode of payment
Payment_installments	Number of installments in case of EMI purchase
Payment_value	Total amount paid for the purchase order

TABLE III. CUSTOMERS DATASET

Column	Description
Customer_id	Customer id
Customer_unique_id	Unique id of the Customer
Customer_zip_code_prefix	Zip code of customer's location
Customer_city	Customer's City
Customer_state	Customer's State

TABLE IV. ORDERS DATASET

Column	Description
Order_id	Order id
Customer_id	Customer id
Order_status	Order status.i.e shipped, delivered
Order_purchase_timeslamp	Timeslamp of the purchase
Order_approved_at	Timeslamp of the order approval

TABLE V. REVIEWS DATASET

Column	Description
review_id	Id of the review given on the product ordered by the order id.
order_id	A unique id of order made by the consumers.
review_score	review score given by the customer for each order on the scale of 1–5.
review_comment_title	Title of the review
review_comment_message	Review comments posted by the consumer for each order.

B. Data Preparation

Once the three sub datasets have been filtered and prepared, the next step is to merge them by their respective primary keys (customer_id, order_id) into a single dataset (Table VI) that will be used to generate other data including the CLV of each customer. However, the generation of CLVs which is done from the models of the Lifetimes library requires, as mentioned above, the application of the RFMTS model with the determination, for each of the 93357client its five variables: recency R, frequency F, monetary value M, regularity of purchase T (Time) and satisfaction S.

C. Generation of Customer Data by Application of the RFMTS Model

At this level, the set of data constituted above (Table VI) will be enriched by the generation for each customer of the five aforementioned factors: Recency, Frequency, Monetary, and Regularity of purchase and Satisfaction value.

To be able to carry out certain calculations, it is necessary to specify the date during which this calculation must be carried out to simulate an immediate study of the transactions of the company. To be consistent with the initial dataset, August 29, 2018, is chosen as the last purchase made. As explained before, recency of each customer is equal to the period between the last purchase and August 29, 2018. Frequency is equal to the total number of purchases made by a customer. Monetary value is equal to the sum of all customer's purchases divided by the total number of purchases. Inter-purchase Time is equal to the period between the last and the first purchase divided by the frequency. The fifth variable, Satisfaction (S), measures the average customer's satisfaction S, it is equal to the sum of all review score divided by the total number of reviews given by a customer.

Table VII presents the new configuration of the study dataset after the generation of R, F, M, T and S for each customer.

D. Clustering Customers based on RFMTS Model

After calculating the variables, Recency, Frequency, Monetary value, Inter-purchase time and satisfaction (Table VII), the next step is grouping the consumers based on the RFMTS model. For this an unsupervised Machine Learning method K-Means is using here.

TABLE VI. INITIAL DATASET

Customer ID	Order ID	Order Purchase Time	Payment Value	Review Score
0a0a…a872	e481f…d6af7	10/2/2017 10:56	10.11	4
c846…84a1	9c5de…053c	10/2/2017 10:56	18.12	4
763c…2b93	11c17…7b62	10/2/2017 10:56	25.59	3
…	…	…	…	…
9b7f…77d6	53cdb2…3451	7/24/2018 20:41	142.14	4
455c…ba1b	4777…ec65d	8/8/2018 8:38	20.46	1

TABLE VII. DATA AFTER CALCULATION F,R,M,T AND S

Customer ID	R	F	M	T	S
0a0a…a872	411.96	2	46.31	121.66	4.34
c846…84a1	140.34	2.17	83.06	121.66	4.77
763c…2b93	210.55	2.30	220.76	91.25	3.17
…	…	…	…	…	…
9b7f…77d6	210.54	1.87	5675.43	170	4.05
455c…ba1b	1350.29	2.15	130.89	320.27	1.24

1) About K-means: As shown in Fig. 2, the k-means algorithm procedure is a simple and easy way to classify a data set through a specified number K of clusters a priori [28], [29]. Each cluster is characterized by its center of gravity, that is, its central point whose coordinates are obtained by calculating the average of each of the coordinates of the sample points assigned to the clusters.

2) K-means customers clustering: Before using it, it was necessary to normalize the variables, because since it is a model that works with the distances of the data, the dimensions and magnitudes play a very important role in its implementation. As an example. Table VIII presents the array of normalized data.

In the case of this study, the "elbow method" was used to indicate the number of clusters which reduces the inertia (proximity of points to their centroid) to a relevant point [30].

Initialization
Choosing the starting points that are used as initial estimates of cluster centroids. They are taken as initial starting values.
Loop:
- **Build k clusters**: examine each point of the data set and assign it to the cluster whose centroid is closest.
- **The new centroids are calculated**: when each point in the data set is assigned to a cluster, it is necessary to recalculate the new centroids k.

Until
- No point changes its cluster assignment or until the centroids no longer move.

Fig. 2. K-means Clustering Algorithm.

TABLE VIII. ARRAY OF NORMALIZED DATA

```
array ([[-0.82886279, -0.84339079, -0.19094101, -0.10203975, -2.29586173],
 [-0.80920189, -0.82379511, -0.19094101, -0.60932832, -0.06528104],
 [ 1.96298496, 1.9391955, -0.19094101, -0.34827659, 0.67824586],
 ...,
 [ 2.16614759, 2.14168417, -0.19094101, -0.23223394, 0.67824586],
 [-0.77643372, -0.79113565, -0.19094101, -0.1383473, 0.67824586],
 [ 1.6156424 , 1.59300518, -0.19094101, -0.41310834, 0.67824586]])
```

TABLE IX. NUMBER OF CUSTOMERS PER CLUSTER

Cluster	Customer
0	14203
1	41730
2	30313
3	5167
4	1994

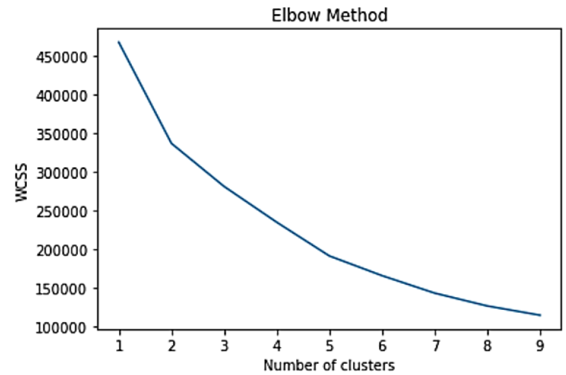

Fig. 3. Elbow Method.

Fig. 3 shows that the number five is the most appropriate for the segmentation. With this number of clusters, the model was trained with the scaled database and attaches the created labels to the initial Table VII.

At first glance, it can be seen from Table IX that the "0" cluster clearly refers to the group with the highest shopping users. A better visualization, of these new segments, can have place by using PCA to reduce the variables into two and be able to translate them into a scatter plot.

K-Means model did a good job of clustering the customers, but from the result of the graph (Fig. 4) and the number of customers per cluster (Table IX), it can be noticed that the order of the clusters does not really represent the importance of the customer versus profitability. This mini theory can be corroborated by looking at the R, F, M, T, S of each cluster (Table X).

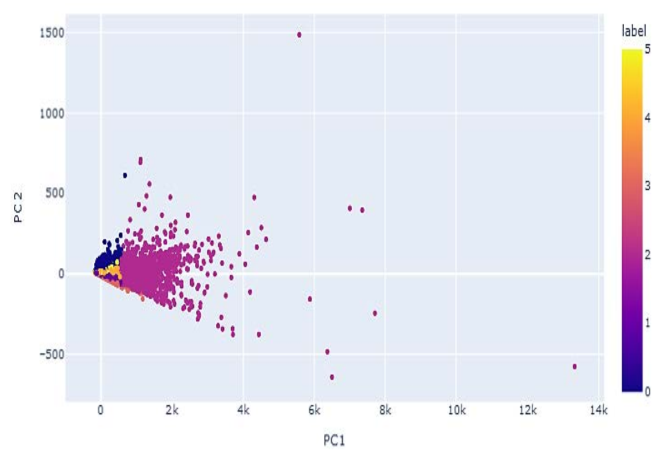

Fig. 4. PCA Dataset Visualization.

TABLE X. AVERAGE RFMTS OF EACH CLUSTER

Cluster	R	F	M	T	S
0	212.91	2.25	204.84	273.75	3.15
1	152.08	2.19	70	30.41	4.65
2	395.41	2	28.33	10.34	4.51
3	1216.66	2.05	100.85	395.41	1.51
4	212.91	1.50	281.91	152.08	4.08

3) Visualisation and analysis of the result: Each cluster represents a segmentation of the market. To make the right decisions and develop efficient marketing strategies, it's necessary to analyze and understand the characteristics of each cluster. The RFMTS characteristics of each cluster are analyzed in this section. Table X lists the average RFMTS values of each cluster. Fig. 5, to Fig. 9 represent the distribution of the five clusters according to the respective distributions R, F, M, T, and S. The x axis of these figures represents the different clusters of customers, while the y axis represents the respective RFMTS values. The height of each rectangle is the range of y value variation, while the y value that lines up exactly with the line inside the rectangle represents the average value. The dataset used in this study contains 93357 customers. Fig. 10 shows the number of customers included in each one of the five clusters (C0, C1, C2, C3, and C4). Among the five clusters, cluster 0 (C0) has the highest percentage of the customers, equal to 44.7%, while cluster 4 (C4) has the lowest percentage of the customers, equal to 2.1%.

Cluster 0: contained 15.2% of the total customers (Fig. 10). It’s characterized by a high R (7.95 months), low F (2.25), high M ($204,84), moderate T (9.46 months) and moderate S (3.15). Most customers in this cluster spend around $204.84. The last shopping date was around 9,46 months ago. The average satisfaction rating is around 3,15. Therefore, this cluster can be labeled as an active satisfied group.

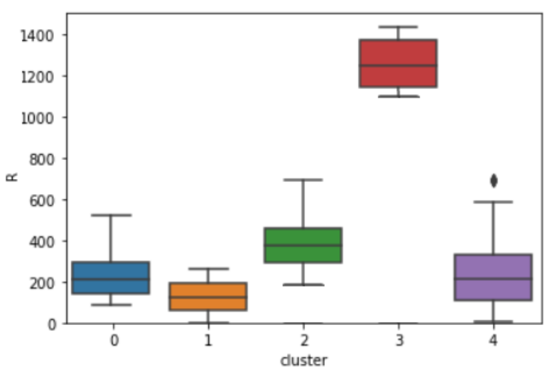

Fig. 5. Distribution of the Five Clusters According to R.

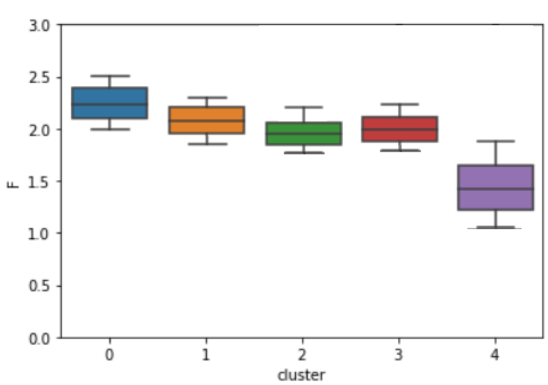

Fig. 6. Distribution of the Five Clusters According to F.

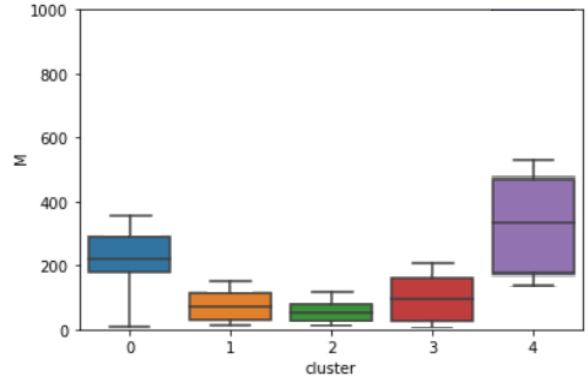

Fig. 7. Distribution of the Five Clusters According to M.

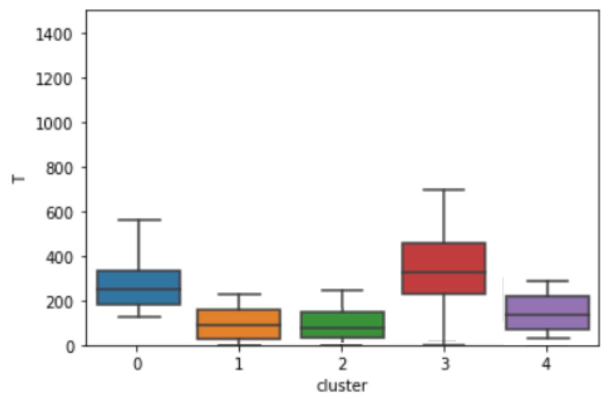

Fig. 8. Distribution of the Five Clusters According to T.

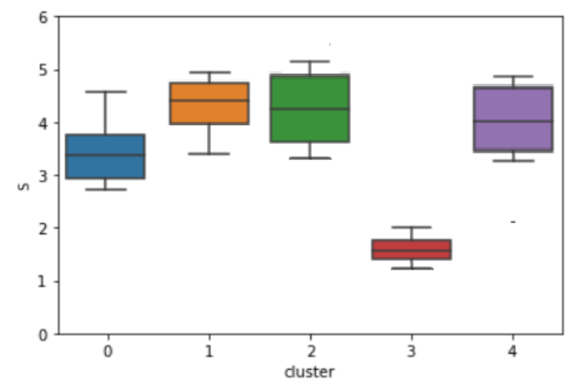

Fig. 9. Distribution of the Five Clusters According to S.

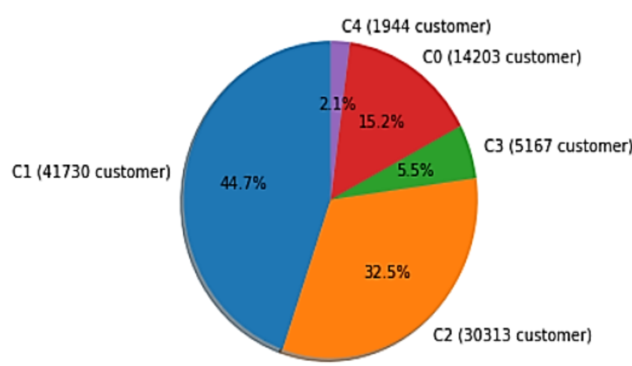

Fig. 10. Clusters Size.

Cluster 1: as the largest sized group, Closter 0 represents 44,7% of all customers (Fig. 10). It's characterized by a high R (5.17 months), low F (2.19), low M ($70), short T (1.17 months) and a high S (4.65). The last purchase date was about 5.17 months ago. Even if M is low, the satisfaction is high, and T is short. Therefore, cluster 0 can be considered as a promising satisfied group.

Cluster 2: represents 32.5% of the total customers (Fig. 10). This cluster is characterized by a low R (13.05 months), low F (2), low M ($28.33), short T (0.34 months) and a high S (4.51). Customers in this cluster spent an average of $28.33, shopped twice in a short time, and their last purchase was around a year ago, and gave high ratings. Therefore, cluster 2 can be considered as the basic group.

Cluster 3: contained 5.5% of the total customers (Fig. 10). It is characterized by a weak R (40.03 months), a low F (2.05), a moderate M ($100.85), a long T (13.05 month) and a low S (1.51). It is the only cluster that simultaneously had low R, F, and S. Even if the average expense is around $100,85, the last purchase in this cluster was about 40,03 months ago, and the satisfaction is really low. Therefore, Cluster 3 can be considered as lost unsatisfied group.

Cluster 4: as the smallest sized group, cluster 4 represents 2,1% of all customers (Fig. 10). This cluster is characterized by a high R (7.56 months), low F (1.50), high M ($281.91), moderate T (5.66) and a high S (4.08). The most recent shopping date was about 7.56 months ago, the average spend is around $281.91, and the ratings are high. For these reasons, this cluster can be labeled as a gold profitable group.

Table XI represent the final table after assigning the corresponding cluster label to each customer.

TABLE XI. FINAL DATASET

Customer ID	R	F	M	T	S	Cluster
0a0a…a872	411.96	2	46.31	121.66	4.34	Basic
c846…84a1	140.34	2.17	83.06	121.66	4.77	Promising satisfied
763c…2b93	210.55	2.30	220.76	91.25	3.17	Active satisfied
…	…	…	…	…	…	…
9b7f…77d6	210.54	1.87	5675.43	170	4.05	Gold profitable
455c…ba1b	1350.29	2.15	130.89	320.27	1.24	Lost unsatisfied

IV. DISCUSSION AND RECOMMENDATIONS

This study is a sort of monograph on the importance of customer satisfaction. Its purpose is to segment customers, the heart of marketing strategy [31]. After pretreating the e-commerce dataset, RFMTS values were calculated for each customer. K-means was then used to segment the customers into five clusters. Finally, a cluster analysis was conducted to identify the characteristics of each customer's cluster. Cluster 0 is regarded as the active satisfied group. C1 is the promising group. While C2 contains basic customers, C3 is thought to be a lost unsatisfied group. C4 is the gold-loyal group. Every company is trying its best to make their customers happy in a business environment that's highly competitive [32], [33]. Compared to previous work based on the RFMT model [1], the clusters obtained (lost group, promising, profitable, active profitable, and loyal) value the profitability of customers but neglect their satisfaction, which is a key factor for any successful and sustainable business. In turbulent markets, a company's reputation largely depends on customer satisfaction [34]. Failing to take customer satisfaction into account can lead to a marketing strategy that is not in line with the interests of customers and therefore have a negative effect on the business. Even if, in the case of this study, the satisfaction rate is high, the company should not rest on its laurels because everything can change, since trade is an area in perpetual transformation (competition, influence, exchange rate).

Finding new customers for the enterprise is essential, but retaining existing customers [35] can be even more important and profitable. Therefore, the company should develop marketing strategies to retain its customers.

- Increate the trust: trust is the belief that the other party of the exchange will not take advantage of opportunistic behavior and will act in a good manner [36]. It has been proven that trust has an important impact on customer purchase intention [37] and that it's a critical factor for online enterprises' success. Keeping the promises made to customers is essential to increasing customer trust and loyalty.
- Increase the net benefits: net benefit of online shopping includes both utilitarian and hedonic value [38]. It's the global benefits received from customers while shopping online. It includes enjoyment, monetary gain, delivery time and after sales service.
- Improve system quality: System quality has a significant impact on online customer satisfaction and decision-making [39], [40]. System quality includes the elements that evaluate the performance of a website, website design, page loading speed, website crashing, interruption, flexibility, and ease of navigation between pages.
- Increase customer reacquisition: customer defection is the decision to end a contract with a specific company [5], [41], [42]. Customer reacquisition provides companies with high economic benefits. An average business has a 60-70% chance of successfully reselling to active buyers, 20-40% to lost customers, and only a 5-20% chance of successfully sale to new prospects [43].

V. CONCLUSION AND OUTLOOK

Segmentation helps to identify potential customers and understand their needs, which increases business revenue. Based on the existent RFMT (i.e., Recency, Frequency, Monetary, and Time) model, satisfaction (S) value was added to create a new model called RFMTS for better customer segmentation. The new model created can be applied to any dataset from which the 5 variables (recency, frequency, monetary, time, and satisfaction) can be extracted. This study started with a typical relational database of an e-commerce site and ended with a segmentation of each of its customers according to their purchasing behaviors and satisfaction. After the customers' clusters are defined, business strategies can be generated to improve customer relationships.

Grouping customers into different groups based of the proposed RFMTS model helps decision makers identify market segments more clearly and thus develop more effective marketing and sales strategies to build customer loyalty [44]. This study opens the door to many outlooks, for example: Reinforced learning can be based on the results obtained in order to identify the cluster target of each type of offers. Being able to identify the best target for a specific offer allow the company to win at several levels, whether it be at the investment level, that the level of customer satisfaction, knowing that they prefer to receive that the adequate offers to their preferences.

ACKNOWLEDGMENT

This research is supported by the Ministry of Higher Education, Scientific Research and Innovation, the Digital Development Agency (DDA) and the National Center for Scientific and Technical Research (CNRST) of Morocco (Smart DLSP Project - AL KHAWARIZMI IA-PROGRAM).

REFERENCES

[1] Z. Jinfeng, W. Jinliang, and X. Bugao, “Customer segmentation by web content mining,” Journal of Retailing and Consumer Services, vol. 61, p. 102588, Jul. 2021, doi: 10.1016/j.jretconser.2021.102588.

[2] S. Gupta et al., “Modeling Customer Lifetime Value,” Journal of Service Research, vol. 9, no. 2, pp. 139–155, Nov. 2006, doi: 10.1177/1094670506293810.

[3] K. P. Sinaga and M.-S. Yang, “Unsupervised K-Means Clustering Algorithm,” IEEE Access, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.

[4] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics, vol. 5.1, pp. 281–298, Jan. 1967.

[5] M. Chandar, A. Laha, and P. Krishna, “Modeling churn behavior of bank customers using predictive data mining techniques.,” 2006, pp. 24–26.

[6] R. Lefébure and G. Venturi, Gestion de la relation client: Edition 2005. Eyrolles, 2011.

[7] P. Padma and J. Ahn, “Guest satisfaction & dissatisfaction in luxury hotels: An application of big data,” International Journal of Hospitality Management, vol. 84, p. 102318, Jan. 2020, doi: 10.1016/j.ijhm.2019.102318.

[8] D. Gursoy and C. Chi, “Effects of COVID-19 pandemic on hospitality industry: review of the current situations and a research agenda,” 2020.

[9] R. Dabestani, A. Shahin, and M. Saljoughian, “Evaluation and prioritization of service quality dimensions based on gap analysis with analytic network process,” 2017, doi: 10.1108/IJQRM-04-2015-0050.

[10] M. A. Shbool, A. Al-Bazi, and R. Al-Hadeethi, “The effect of customer satisfaction on parcel delivery operations using autonomous vehicles: An agent-based simulation study,” Heliyon, vol. 8, no. 5, p. e09409, May 2022, doi: 10.1016/j.heliyon.2022.e09409.

[11] A. M. Hughes, Strategic database marketing: the masterplan for starting and managing a profitable, customer-based marketing program. Chicago, Ill: Probus Pub. Co., 1994.

[12] I. Maryani, D. Riana, R. D. Astuti, A. Ishaq, Sutrisno, and E. A. Pratama, “Customer Segmentation based on RFM model and Clustering Techniques With K-Means Algorithm,” in 2018 Third International Conference on Informatics and Computing (ICIC), Oct. 2018, pp. 1–6. doi: 10.1109/IAC.2018.8780570.

[13] A. J. Christy, A. Umamakeswari, L. Priyatharsini, and A. Neyaa, “RFM ranking - An effective approach to customer segmentation,” J. King Saud Univ. Comput. Inf. Sci., 2018, doi: 10.1016/J.JKSUCI.2018.09.004.

[14] I.-C. Yeh, K.-J. Yang, and T.-M. Ting, “Knowledge discovery on RFM model using Bernoulli sequence,” Expert Systems with Applications, vol. 36, no. 3, Part 2, pp. 5866–5871, Apr. 2009, doi: 10.1016/j.eswa.2008.07.018.

[15] M. Alvandi, S. Fazli, and F. Abdoli, “K-Mean Clustering Method For Analysis Customer Lifetime Value With LRFM Relationship Model In Banking Services,” 2012. https://www.semanticscholar.org/paper/K-Mean-Clustering-Method-For-Analysis-Customer-With-Alvandi-Fazli/54ac58708a608bee195e9f60825a61da51ebbcf1 (accessed May 17, 2022).

[16] S. Peker, A. Kocyigit, and P. E. Eren, “LRFMP model for customer segmentation in the grocery retail industry: a case study,” Marketing Intelligence & Planning, vol. 35, no. 4, pp. 544–559, Jan. 2017, doi: 10.1108/MIP-11-2016-0210.

[17] S. Q. Moghaddam, N. Abdolvand, and S. Rajaee Harandi, “A RFMV model and customer segmentation based on variety of products,” Journal of Information Systems and Telecommunication, vol. 5, pp. 155–161, Jun. 2017.

[18] F. Yoseph and M. Heikkila, “Segmenting Retail Customers with an Enhanced RFM and a Hybrid Regression/Clustering Method,” 2018 International Conference on Machine Learning and Data Engineering (iCMLDE), 2018, doi: 10.1109/ICMLDE.2018.00029.

[19] A. P. Darko and D. Liang, “Modeling customer satisfaction through online reviews: A FlowSort group decision model under probabilistic linguistic settings,” Expert Systems with Applications, vol. 195, p. 116649, Jun. 2022, doi: 10.1016/j.eswa.2022.116649.

[20] A. Ahani et al., “Revealing customers’ satisfaction and preferences through online review analysis: The case of Canary Islands hotels,” Journal of Retailing and Consumer Services, vol. 51, pp. 331–343, Nov. 2019, doi: 10.1016/j.jretconser.2019.06.014.

[21] J.-W. Bi, Y. Liu, Z.-P. Fan, and E. Cambria, “Modelling customer satisfaction from online reviews using ensemble neural network and effect-based Kano model,” International Journal of Production Research, vol. 57, no. 22, pp. 7068–7088, Nov. 2019, doi: 10.1080/00207543.2019.1574989.

[22] C. Fornell, “A National Customer Satisfaction Barometer: The Swedish Experience,” Journal of Marketing, vol. 56, no. 1, pp. 6–21, 1992, doi: 10.2307/1252129.

[23] R. L. Oliver, Satisfaction: A Behavioral Perspective on the Consumer. McGraw Hill, 1997.

[24] D. Lee, J. Moon, Y. J. Kim, and M. Y. Yi, “Antecedents and consequences of mobile phone usability: Linking simplicity and interactivity to satisfaction, trust, and brand loyalty,” Information & Management, vol. 52, no. 3, pp. 295–304, Apr. 2015, doi: 10.1016/j.im.2014.12.001.

[25] C. Liao, H.-N. Lin, M. M. Luo, and S. Chea, “Factors influencing online shoppers’ repurchase intentions: The roles of satisfaction and regret,” Inf. Manag., 2017, doi: 10.1016/j.im.2016.12.005.

[26] E. Park, Y. Jang, J. Kim, N. J. Jeong, K. Bae, and A. P. del Pobil, “Determinants of customer satisfaction with airline services: An analysis of customer feedback big data,” Journal of Retailing and Consumer Services, vol. 51, pp. 186–190, Nov. 2019, doi: 10.1016/j.jretconser.2019.06.009.

[27] Y. Piris and A.-C. Gay, “Customer satisfaction and natural language processing,” Journal of Business Research, vol. 124, pp. 264–271, Jan. 2021, doi: 10.1016/j.jbusres.2020.11.065.

[28] A. Banumathi and A. Pethalakshmi, “Refinement of K-Means and Fuzzy C-Means,” International Journal of Computer Applications, vol. 39, no. 17, Feb. 2012, Accessed: Oct. 21, 2021. [Online]. Available: https://www.ijcaonline.org/archives/volume39/number17/4911-7441.

[29] R. J. Bharati and V. V. Gohokar, "A Comparative Analysis of Fuzzy C-Means Clustering and K Means Clustering Algorithms," International Journal Of Computational Engineering Research, Jun. 2012.

[30] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, "Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster," IOP Conf. Ser.: Mater. Sci. Eng., vol. 336, p. 012017, Apr. 2018, doi: 10.1088/1757-899X/336/1/012017.

[31] M. S. Kahreh, M. Tive, A. Babania, and M. Hesan, "Analyzing the Applications of Customer Lifetime Value (CLV) based on Benefit Segmentation for the Banking Sector," Procedia - Social and Behavioral Sciences, vol. 109, pp. 590–594, Jan. 2014, doi: 10.1016/j.sbspro.2013.12.511.

[32] K. Hellen and M. Saaksjarvi, "Happiness as a predictor of service quality and commitment for utilitarian and hedonic services," Psychology & Marketing, vol. 28, no. 9, pp. 934–957, 2011.

[33] A. Kumar, J. Paul, and S. Starčević, "Do brands make consumers happy?- A masstige theory perspective," Journal of Retailing and Consumer Services, vol. 58, p. 102318, Jan. 2021, doi: 10.1016/j.jretconser.2020.102318.

[34] R. U. Khan, Y. Salamzadeh, Q. Iqbal, and S. Yang, "The Impact of Customer Relationship Management and Company Reputation on Customer Loyalty: The Mediating Role of Customer Satisfaction," Journal of Relationship Marketing, vol. 21, no. 1, pp. 1–26, Jan. 2022, doi: 10.1080/15332667.2020.1840904.

[35] L. Tong, Y. Wang, F. Wen, and X. Li, "The research of customer loyalty improvement in telecom industry based on NPS data mining," China Communications, vol. 14, no. 11, pp. 260–268, Nov. 2017, doi: 10.1109/CC.2017.8233665.

[36] D. Gefen, E. Karahanna, and D. Straub, "Trust and TAM in Online Shopping: An Integrated Model," MIS Q., 2003, doi: 10.2307/30036519.

[37] M. Tajvidi, Y. Wang, N. Hajli, and P. E. D. Love, "Brand value Co-creation in social commerce: The role of interactivity, social support, and relationship quality," Computers in Human Behavior, vol. 115, p. 105238, Feb. 2021, doi: 10.1016/j.chb.2017.11.006.

[38] T. L. Childers, C. L. Carr, J. Peck, and S. Carson, "Hedonic and utilitarian motivations for online retail shopping behavior," Journal of Retailing, vol. 77, no. 4, pp. 511–535, Dec. 2001, doi: 10.1016/S0022-4359(01)00056-2.

[39] H. H. Bauer, T. Falk, and M. Hammerschmidt, "eTransQual: A transaction process-based approach for capturing service quality in online shopping," Journal of Business Research, vol. 59, no. 7, pp. 866–875, Jul. 2006, doi: 10.1016/j.jbusres.2006.01.021.

[40] D. M. Szymanski and R. T. Hise, "E-satisfaction: an initial examination," Journal of Retailing, vol. 76, no. 3, pp. 309–322, Jul. 2000, doi: 10.1016/S0022-4359(00)00035-X.

[41] K. Stewart, "An exploration of customer exit in retail banking," International Journal of Bank Marketing, vol. 16, no. 1, pp. 6–14, Jan. 1998, doi: 10.1108/02652329810197735.

[42] G. Capponi, N. Corrocher, and L. Zirulia, "Personalized pricing for customer retention: Theory and evidence from mobile communication," Telecommunications Policy, vol. 45, no. 1, p. 102069, Feb. 2021, doi: 10.1016/j.telpol.2020.102069.

[43] J. Griffin and M. W. Lowenstein, Customer Winback: How to Recapture Lost Customers--And Keep Them Loyal. Wiley, 2001.

[44] A. Hiziroglu and S. Sengul, "Investigating Two Customer Lifetime Value Models from Segmentation Perspective," Procedia - Social and Behavioral Sciences, vol. 62, pp. 766–774, Oct. 2012, doi: 10.1016/j.sbspro.2012.09.129.