International Journal of Technology and Engineering Studies
volume 7 issue 2 pp. 8-14 doi: https://dx.doi.org/10.20469/ijtes.7.10002-2

# Fuzzy RFM Analysis in Car Rental Sector

**Onur Dogan** *
Department of Industrial Engineering,
Izmir Bakircay University,
Izmir, Turkey

**Basar Oztaysi**
Department of Industrial Engineering,
Istanbul Technical University,
Istanbul, Turkey

**Aydeniz Isik**
Ataturk Mah,
Atasehir Blv. 49 Ada Kamelya,
Istanbul, Turkey

***Abstract:*** Recency, Frequency, Monetary (RFM) technique is a common and useful way for customer segmentation analysis, which is a must in sales, marketing, and operation management. RFM mainly uses transaction data to investigate customer's shopping behaviour. This study applied a fuzzy-based RFM method that uses renting data from a car rental company. First, data were extracted from the database and were transformed to R, F, and M parameters. Second, R, F, and M parameters were normalized and converted to fuzzy numbers. A fuzzy c - means FCM clustering algorithm was applied to transform fuzzy R, F, and M numbers to make groups for customers. As a result, some customers were regarded as regular customers, where as others could be divided into wintertime and summertime customers. Managers in the company could make better decisions and offer more relevant promotions for specific customer groups.

***Keywords:*** *Customer segmentation, fuzzy RFM, fuzzy clustering, shopping behaviour, car rental*

***Received:*** 16 June 2021; ***Accepted:*** 19 July 2021; ***Published:*** 24 August 2021

## I. INTRODUCTION

Globalization has a profound impact on competition for companies. Today's technological developments enable companies to improve their marketing strategy, promotions, or campaigns special to customer groups. Customer segmentation is a fundamental approach to reach success in these kinds of actions. Customer segmentation studies focus on customer characteristics such as age, gender, location, income to reveal shopping behaviour. Companies use customer segmentation results to make business decisions about marketing tactics and strategies by identifying customers' trends.

RFM is a practical analysis used to segment customers. It is used historical transaction data to assess customer's shopping behaviour. Recency indicates how recently a customer has purchased among the three main dimensions. Frequency shows the customer's purchasing frequency, whereas monetary states how much the customer spends. RFM is frequently utilized to measure customer value considering the past purchasing records [1]. With a scoring method, the approach computes R, F, and M scores combined in the groups generated to understand customers' shopping behavior. As a result of RFM, customers with low recency, high frequency, and high monetary values are described as significant customers [2, 3]. Companies can predict future purchase promotions to increase the satisfaction of significant customers and keep customers who have a possible churn risk.

Soft computing methods have been used in customer segmentation research as data-based methods. Soft com-

*Correspondence concerning this article should be addressed to Onur Dogan, Department of Industrial Engineering, Izmir Bakircay University, Izmir, Turkey. E-mail: onur.dogan@bakircay.edu.tr

puting approaches shape future segmentation studies because of the special applications [4, 5, 6]. As a soft computing method, fuzzy-based clustering methods are considered the most notable methods [7, 4]. Recent literature includes different versions of the fuzzy clustering methods employed in various domains such as segmentation [8, 9] computer vision [10], customer path segmentation [4]. This study adopts the FCM technique, one of the most popular fuzzy-based clustering methods.

The remaining part of the study is organized as follows. The study first represents the applied methodology, including RFM and FCM. Section 3 gives the results of the proposed method. Section 4 concludes the study and discusses the results.

## II. LITERATURE REVIEW

Various customer data have been used to perform customer segmentation problems in the literature. Therefore, the proposed segmentation techniques are diverse, such as lifetime period, demographic-based, propensity-based, and value-based segmentation [11]. When we look at the methods used for segmentation, it can be seen that they can rely on basic arithmetic operations, expert experiences, or analytical methods. Traditional RFM analysis uses basic arithmetic operations, whereas cluster analysis is a data mining technique that is highly adopted for analytic segmentation [12].

Many RFM studies have been combined with other techniques in the literature to increase the effectiveness of customer segmentation. [13] combined the RFM and rough set theory to implement a value-based segmentation. [14] integrated RFM and sequential pattern mining by using the customers' buying data. [15] introduced a novel customer segmentation methodology, including RFM ranking. [16] combined the RFM and k-means clustering technique to segment customers considering their purchasing data. [17] applied RFM analysis to predict the customer value for each product [18] found the customers' lifetime value using RFM analysis. Some studies focused on comparative research. [19] and Hastak (2007) and [20] compared RFM with basic segmentation methods such as decision tree and logistic regression. [21] applied customer churn analysis using RFM approach and compared various deep learning and machine learning methods.

Recent investigations on fuzzy customer segmentation have focused on various data sources such as location data [22, 23] and social media data. [22] used location data using a beacon network and explained a novel customer segmentation method. [4] proposed a segmentation method considering the number of visiting a shopping mall. They created customers' routes from location data and applied FCM clustering algorithm. [24] used indoor activities of customers. The authors clustered the customers' paths and predicted customers' gender information by using fuzzy c-medoids clustering. [25] collected social media data belonging to SPA hotels to introduce a decision-making prediction model. [26] used Twitter data and applied the FCM method to make relevant groups of Twitter users.

## III. METHODOLOGY

### A. *RFM Analysis*

RFM analysis, one of the well-known segmentation approaches, is based on three purchasing attributes, which are recency (R) of the purchase, frequency (F) of the purchase, and monetary (M) value of the purchase. Recency indicates the time between the recent shopping date and the present. Marketers think that many recent customers are more likely to buy again than less-recent customers. Frequency shows how frequently a customer purchases within a certain period. It is assumed that the higher frequency, the more purchasing activity. Monetary refers to the total amount of money spent by the customer.

Traditional RFM analysis starts with calculating the recency, frequency, and monetary value, respectively. Customers are sorted in descending order considering recency value in the first step. The customers are then split into a predefined number of groups. Assuming the number is selected as 3, the data in each perspective is sorted, and the first 33% is coded as 1, the next 33% is coded as 2 and the final 33% is coded as 3. Since there are three perspectives, each perspective is divided into three groups. For each customer R, F, and M codes are calculated and then combined. It is possible to use an analytic network process with RFM. For example, [27] propose weighted RFM in which R, F, M values are multiplied by the relative weights. Segmentation is executed using these weighted values. [28] proposed timely RFM considering product sales at different times.

### B. *FCM Clustering*

As one of the most popular clustering methods, FCM is a soft version of k-means clustering. It considers that one transaction of a customer can belong to more than one cluster with various membership degrees. In real-life cases, data points may have nearly equal similarities to more than one cluster, called boundary data points. Fuzzy clustering methods provide better accuracies by assigning the boundary data points into the proper segment [24]. The FCM attempts to minimize Eq 1.

$$J(U,V:X)=\sum_{k=1}^{c}\sum_{i=1}^{N}\mu_{ki}^{m}d^{2}\left(v_{k},x_{i}\right) \quad (1)$$

Where U is the fuzzy membership matrix, V is the cluster centers, and X is the data set to be clustered. $\mu_k i$ is the membership value of data $x_i$ fuzzy cluster $c_k$. The fuzzifier m must be larger than 1. If m equals to 1, then the clusters are formed in a crisp format. An appropriate U and V must be computed like Eq 2 and Eq 3 to minimize J, for all k = 1,2,c.

$$\mu_{ki}=\left(\sum_{j=1}^{c}\left(d_{ki}^{2}/d_{ji}^{2}\right)^{1/(m-1)}\right)^{-1} \quad (2)$$

$$v_{k}=\frac{\sum_{i=1}^{N}\mu_{ki}^{m}x_{i}}{\sum_{i=1}^{N}\mu_{ki}^{m}} \quad (3)$$

Eq 2 and Eq 3 are used until the termination criterion is satisfied.

### C. *Fuzzy RFM Analysis in Car Rental Industry*

A car rental agency is a company that rents automobiles for short periods to the public, generally ranging from a few hours to a few weeks. Car rental brands are generally structured with numerous local branches, which allow users to return a vehicle to a different location and are primarily located near airports. The reservations can be complemented by using a website or call center.

In this case study, car rental data provided customer segmentation by using fuzzy RFM. At the first stage of the study, the transactional data were prepared for RFM analysis. To this end, a data set containing Customer Id, Customers average basket size, Customer's last transaction date, and the number of transactions in the last year were prepared. The sample data set is given in Table 1.

TABLE 1
THE AGGREGATED DATA FOR FUZZY RFM

| Customer Id | Number of Days from Last Order (R) | Number of Orders in the Period (F) | Average Basket Size (M) |
|---|---|---|---|
| 10001 | 45 | 3 | 250 |
| .... | | | |
| n | 40 | 4 | 140 |

Using the same procedure, three datasets are prepared, namely, RFM - Winter, RFM - Summer, RFM - Year. As the name implies, the RFM - Winter dataset represents the table considering the winter transactions only, the RFM - Summer dataset represents the summer transactions only. Finally, RFM - Year represents the transactions during the last year.

Fuzzy clustering was applied to each data set for different cluster numbers starting from three to seven. To represent the performance of clustering, Explained Variance was calculated. Fig. 1 and Table 2 show the clustering results and associated explained variance values.

TABLE 2
EXPLAIN VARIANCE VALUES FOR FUZZY CLUSTERING APPLICATIONS

| | c = 3 | c = 4 | c = 5 | c = 6 | c = 7 |
|---|---|---|---|---|---|
| RFM - Year | 0.725 | 0.729 | 0.706 | 0.741 | 0.741 |
| RFM - Winter | 0.845 | 0.889 | 0.895 | 0.924 | 0.932 |
| RFM - Summer | 0.765 | 0.838 | 0.865 | 0.883 | 0.895 |

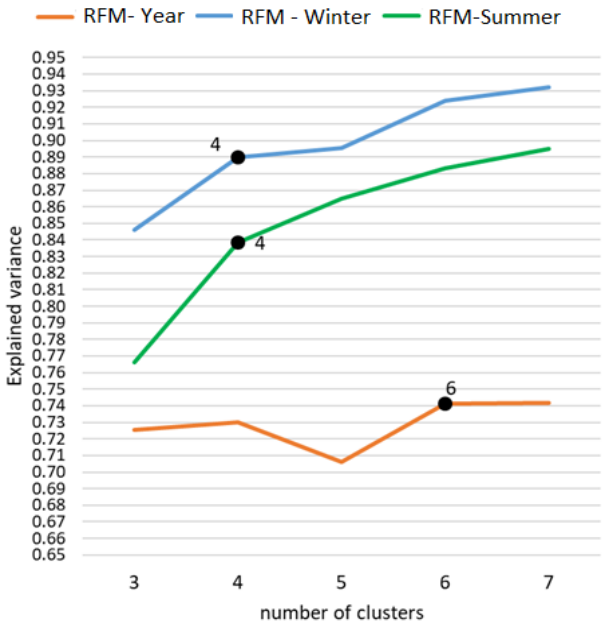


Fig. 1. Fuzzy clustering results and explained variance

By using the elbow technique, the optimal cluster size for RFM - Year was selected as 6, and for others, the optimum value was determined as 4. The results revealed that the explained variance decreases when yearly data is used. The centroid table of RFM - Year was analyzed with the experts, and it was observed that the resulting clusters were not actionable. (Due to page restrictions, the Centroid table for RFM - Year was not provided here). To get more actionable results, the centroid tables of RFM - Winter and RFM - Summer were analyzed. At the first step, the centroid table and cluster sizes for RFM - Winter and RFM - Summer were evaluated separately. The cluster number is determined as four in both datasets, but the experts had difficulties explaining all of the clusters.

For both cases, three segments can be defined as; Platinum: highly active and profitable customers, Standard: Customers with moderate activity, and Passive: The customers who are not recently active. Indeed, numeric results, given in Table 3, provide clusters entitled Standard-Low and Standard-High, but the experts guided the study to combine these two clusters.

TABLE 3
CENTROID TABLES FOR RFM - WINTER AND RFM - SUMMER

| Centroids | C 1 | C 2 | C 3 | C 4 |
|---|---|---|---|---|
| Winter - R | -1.0105 | -1.0104 | 0.8992 | -1.0256 |
| Winter - F | -1.1335 | -0.9590 | 0.8961 | -0.9411 |
| Winter - M | 1.2504 | 0.2543 | -0.4410 | 0.1929 |
| Member size | 465 | 353 | 2178 | 1155 |
| Centroids | C 1 | C 2 | C 3 | C 4 |
| Summer - R | -0.7119 | 1.2793 | -0.7552 | -0.7162 |
| Summer - F | -0.6047 | 1.2336 | -0.8492 | -0.6424 |
| Summer - M | 0.0581 | -0.6040 | 0.9450 | 0.1656 |
| Member size | 1476 | 1474 | 695 | 506 |

Table 3 shows the cluster centroids by using the normalized values. As explained before, C 2 and C 4 in RFM - Winter and C 1 and C 4 in RFM-Summer have similar values, so they were combined to form a new segment.

The results of RFM-Winter and RFM - Summer provide helpful information for their period, but they do not carry information from an annual perspective. To overcome this issue, the segmentation information of each customer was combined as given in Table 4. The first raw in Table 4 shows that Customer 48 is a member of Passive Segment during winter, but it is a member of Standard Segment during summer. However, Customer 54 is a member of the Platinum Segment both in winter and summer.

TABLE 4
SAMPLE COMBINED CUSTOMER SEGMENTS

| Customer Id | RFM - Winter | RFM - Summer |
|---|---|---|
| 48 | Passive | Normal |
| 54 | Platinum | Platinum |
| .. | .. | .. |
| 88 | Passive | Platinum |

As a result of the study, eight combined segments were formed in Table 5. The results revealed that 66 customers are in the Platinum-Platinum segment, which means they are the most active and profitable customers during the year. On the other hand, 1652 customers are in the Passive-Standard segment, which means they do not rent cars during winter and rarely rent cars during Summer. The number of cluster members is reasonable since no single cluster has a high portion of whole customers.

TABLE 5
COMBINED SEGMENTS AND NUMBER OF MEMBERS

| Combined Segments (Winter - Summer) | Number of Members |
|---|---|
| Passive - Standard | 1652 |
| Passive - Platinum | 526 |
| Standard - Passive | 1144 |
| Standard - Standard | 264 |
| Standard - Platinum | 100 |
| Platinum - Passive | 330 |
| Platinum - Normal | 66 |
| Platinum - Platinum | 69 |

## IV. CONCLUSION

This study focused on customer segmentation in the car rental industry. Using a sample of real-world data, the fuzzy ARM analysis was applied. The contribution of the paper can be given in two parts. First, it uses FCM algorithm for RFM analysis, and second, it accomplished segmentation for different periods and then created a combined segmentation. This approach is beneficial since it can show customer behaviours in different periods.

## V. ACKNOWLEDGMENT

This work is supported by TUBITAK (The Scientific and Technological Research Council of Turkey, Project Id: 3200643)

## REFERENCES

[1] Y.-H. Hu and T.-W. Yeh, "Discovering valuable frequent patterns based on RFM analysis without customer identification information," *Knowledge-Based Systems*, vol. 61, pp. 76–88, 2014. doi: https://doi.org/10.1016/j.knosys.2014.02.009

[2] A. N. Noorzad and T. Sato, "Multi-criteria fuzzy-based handover decision system for heterogeneous wireless networks,," *International Journal of Technology and Engineering Studies*, vol. 3, no. 4, pp. 159–168, 2017.

[3] M. A. Rahim, M. Mushafiq, S. Khan, and Z. A. Arain, "RFM - based repurchase behavior for customer classification and segmentation," *Journal of Retailing and Consumer Services*, vol. 61, pp. 1–9, 2021. doi: https://doi.org/10.1016/j.jretconser.2021.102566

[4] O. Dogan, "Segmentation analysis of companies natural gas consumption by soft clustering," in *International Conference on Intelligent and Fuzzy Systems:* Cham, UK, 2020. doi: https://doi.org/10.1007/978-3-030-51156-2_7

[5] A. Hiziroglu, "A neuro-fuzzy two-stage clustering approach to customer segmentation," *Journal of*

*Marketing Analytics*, vol. 1, no. 4, pp. 202–221, 2013. doi: https://doi.org/10.1057/jma.2013.17

[6] H. Zhang, G. Kou, and Y. Peng, "Soft consensus cost models for group decision making and economic interpretations," *European Journal of Operational Research*, vol. 277, no. 3, pp. 964–980, 2019. doi: https://doi.org/10.1016/j.ejor.2019.03.009

[7] O. Dogan and B. Oztaysi, "From indoor paths to gender prediction with soft clustering," *Journal of Intelligent & Fuzzy Systems*, vol. 39, no. 5, pp. 6529–6538, 2020. doi: https://doi.org/10.3233/JIFS-189116

[8] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognition*, vol. 40, no. 3, pp. 825–838, 2007. doi: https://doi.org/10.1016/j.patcog.2006.07.011

[9] Y. Yong, Z. Chongxun, and L. Pan, "A novel fuzzy c-means clustering algorithm for image thresholding," *Measurement Science Review*, vol. 4, no. 1, pp. 11–19, 2004.

[10] D. K. Iakovidis, N. Pelekis, E. Kotsifakos, and I. Kopanakis, "Intuitionistic fuzzy clustering with applications in computer vision," in *International Conference on Advanced Concepts for Intelligent Vision Systems:* Berlin, Heidelberg, 2008. doi: https://doi.org/10.1007/978-3-540-88458-3_69

[11] K. K. Tsiptsis and A. Chorianopoulos, *Data mining techniques in CRM: Inside customer segmentation*. Hoboken, NJ: John Wiley & Sons, 2011.

[12] J. Han, J. Pei, and M. Kamber, *Data mining: Concepts and techniques*. *Amsterdam, Netherlands*: Elsevier, 2011.

[13] C.-H. Cheng and Y.-S. Chen, "Classifying the segmentation of customer value via RFM model and RS theory," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4176–4184, 2009. doi: https://doi.org/10.1016/j.eswa.2008.04.003

[14] Y.-L. Chen, M.-H. Kuo, S.-Y. Wu, and K. Tang, "Discovering Recency, Frequency, and Monetary RFM sequential patterns from customers purchasing data," *Electronic Commerce Research and Applications*, vol. 8, no. 5, pp. 241–251, 2009. doi: https://doi.org/10.1016/j.elerap.2009.03.002

[15] A. J. Christy, A. Umamakeswari, L. Priyatharsini, and A. Neyaa, "RFM ranking-an effective approach to customer segmentation," *Journal of King Saud University-Computer and Information Sciences*, vol. 33, no. 10, pp. 1251–1257, 2021. doi: https://doi.org/10.1016/j.jksuci.2018.09.004

[16] P. Anitha and M. M. Patil, "RFM model for customer purchase behavior using k-means algorithm (just accepted)," *Journal of King Saud University-Computer and Information Sciences*, 2019. doi: https://doi.org/10.1016/j.jksuci.2019.12.011

[17] R. Heldt, C. S. Silveira, and F. B. Luce, "Predicting customer value per product: From RFM to RFM/p," *Journal of Business Research*, vol. 127, pp. 444–453, 2021. doi: https://doi.org/10.1016/j.jbusres.2019.05.001

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

[19] J. A. McCarty and M. Hastak, "Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression," *Journal of Business Research*, vol. 60, no. 6, pp. 656–662, 2007. doi: https://doi.org/10.1016/j.jbusres.2006.06.015

[20] K. Coussement, F. A. Van den Bossche, and K. W. De Bock, "Data accuracy's impact on segmentation performance: Benchmarking RFM analysis, logistic regression, and decision trees," *Journal of Business Research*, vol. 67, no. 1, pp. 2751–2758, 2014. doi: https://doi.org/10.1016/j.jksuci.2018.09.004

[21] O. F. Seymen, O. Dogan, and A. Hiziroglu, "Customer churn prediction using deep learning," in *International Conference on Soft Computing and Pattern Recognition:* Cham, UK, 2020.

[22] S. C. Oner and B. Oztaysi, "An interval type 2 hesitant fuzzy mcdm approach and a fuzzy c means clustering for retailer clustering," *Soft Computing*, vol. 22, no. 15, pp. 4971–4987, 2018. doi: https://doi.org/10.1007/s00500-018-3191-0

[23] S. C. Öner and B. Öztayşi, "An interval valued hesitant fuzzy clustering approach for location clustering and customer segmentation," in *Advances in Fuzzy Logic and Technology*. Cham, UK: Springer, 2017. doi: https://doi.org/10.1007/978-3-319-66827-7_6

[24] O. Dogan and B. Oztaysi, "Gender prediction from classified indoor customer paths by fuzzy c-medoids clustering," in *International Conference on Intelligent and Fuzzy Systems:* Cham, UK, 2019. doi: https://doi.org/10.1007/978-3-030-23756-1_21

[25] A. Ahani, M. Nilashi, O. Ibrahim, L. Sanzogni, and S. Weaven, "Market segmentation and travel choice prediction in spa hotels through tripadvisors online reviews," *International Journal of Hos-*

*pitality Management*, vol. 80, pp. 52–77, 2019. doi: https://doi.org/10.1016/j.ijhm.2019.01.003

[26] B. Öztaysi and S. Ç. Onar, "User segmentation based on twitter data using fuzzy clustering," in *Data Mining in Dynamic Social Networks and Fuzzy Systems*. *PA, United States*: IGI Global, 2013.

[27] A. Zare Ravasan and T. Mansouri, "A dynamic ERP critical failure factors modelling with FCM throughout project lifecycle phases," *Production Planning & Control*, vol. 27, no. 2, pp. 65–82, 2016. doi: https://doi.org/10.1080/09537287.2015.1064551

[28] L. Lee and W.-J. Liu, "The timely product recommendation based on RFM method," in *International Conference on Business and Information,* Singapore, 2006.