

RFM vs Clustering: A Comparative Study of Customer Segmentation Approaches

Assistant Professor Smitha Mangesh Junnarkar¹

Mubashira Farid Khan²

¹Department of Data Science

SIES College of Arts, Science and Commerce

(Empowered Autonomous), Sion (W), Mumbai

smitaj@sies.edu.in

²Department of Data Science

SIES College of Arts, Science and Commerce

(Empowered Autonomous), Sion (W), Mumbai

mfk3120@gmail.com

October 17, 2026

Abstract

Customer segmentation practices enable businesses to categorize clients into segments of customers with similar purchasing behaviors. This technique helps the company to develop its marketing strategy, identify key customers and establish ways to build long-term relationships with them, and optimize the marketing spend. This article aims to compare two customer segmentation techniques, namely Recency, Frequency, and Monetary (RFM) analysis and clustering. Whereas the former technique relies on several simple quantitative indicators, the latter technique uses machine learning methods to find intrinsic patterns in data. The current study provides a description of how customers can be segmented using these two approaches, compares their advantages and disadvantages, and explores the cases when one method may be more efficient than the other. The article concludes that RFM analysis is an easy-to-implement solution that allows businesses to obtain a rough understanding of customers' purchasing behaviors, whereas the clustering method is more sophisticated and flexible to reveal hidden patterns in customers' behaviors. These methods can also be combined to analyze the relationship between customers' values, determined by the RFM model, and their clusters.

Keywords: Customer segmentation, RFM analysis, clustering, K-Means, customer relationship management, marketing strategy

1 Introduction

Businesses deal with customers who differ in purchasing habits, spending capacity, loyalty, and product preferences. Treating every customer in the same way may reduce the effectiveness of marketing activities. Customer segmentation solves this problem by dividing customers into meaningful groups so that companies can offer suitable products, promotions, and communication to each group (Alves Gomes and Meisen, 2023).

RFM analysis is one of the most popular traditional methods for customer segmentation. It evaluates customers using three variables: Recency, Frequency, and Monetary value. Recency measures the time since a customer's last purchase, Frequency measures the number of purchases, and Monetary value measures the total amount spent (Christy et al., 2021). This method is widely used because it is simple, easy to interpret, and directly related to customer purchasing behaviour.

Clustering is another important approach used in customer segmentation. Unlike RFM analysis, clustering is an unsupervised machine-learning technique that groups customers based on similarities in their data. K-Means is one of the most commonly used clustering algorithms because it is simple, scalable, and suitable for numerical data (John et al., 2023). Clustering can use many customer variables, such as purchase frequency, spending amount, product preferences, and website activity.

This paper compares RFM analysis and clustering as customer segmentation approaches. The main purpose is not to declare one method better than the other, but to understand their differences and identify when each method is more useful.

2 Research Objectives

The objectives of this research paper are:

1. To explain the basic concepts and working process of RFM analysis and clustering for customer segmentation.
2. To compare RFM analysis and clustering based on simplicity, interpretability, data requirements, and usefulness for marketing decisions.
3. To identify situations where RFM, clustering, or a combined RFM-clustering approach is most suitable for businesses.

3 Literature Review

RFM analysis has been widely used in marketing and customer relationship management. It helps businesses identify valuable customers by examining how recently they purchased, how frequently they purchase, and how much they spend (Christy et al., 2021). Customers with recent purchases, high purchase frequency, and high spending are generally considered more valuable than customers with low activity.

A review of customer segmentation methods in e-commerce found that RFM analysis is one of the most common ways to represent customer behaviour before segmentation (Alves Gomes and Meisen, 2023). The review examined 105 publications published between 2000 and 2022 and found that RFM is often used because it can be easily adapted to different business situations.

Clustering methods have also been widely applied in customer segmentation. K-Means clustering divides customers into a fixed number of groups based on the distance between customer data points and cluster centres (John et al., 2023). It is often combined with RFM values to create customer groups such as loyal customers, potential customers, inactive customers, and high-value customers.

Research on retail customer segmentation has compared several clustering algorithms, including K-Means, Gaussian Mixture Models, DBSCAN, agglomerative clustering, and BIRCH. The study showed that clustering techniques can identify meaningful customer groups when applied to RFM-based customer data (John et al., 2023). However, the performance of clustering depends on data quality, feature selection, scaling, and the choice of the number of clusters.

4 Methodology

This study uses a comparative and conceptual research approach. It reviews existing academic literature on RFM analysis and clustering-based customer segmentation. The comparison is based on four main criteria:

- **Simplicity:** How easy the method is to understand and implement.
- **Interpretability:** How clearly the resulting customer groups can be explained.
- **Data requirements:** The type and amount of data needed.
- **Marketing usefulness:** How well the results support marketing decisions.

The study mainly focuses on K-Means clustering because it is one of the most widely used clustering algorithms in customer segmentation research (Alves Gomes and Meisen, 2023; John et al., 2023).

5 RFM Analysis

5.1 Concept of RFM

RFM analysis is a type of marketing procedure that is usually used to group the customers and quantitatively rank the customers based on the recency, frequency, and monetary values of their transactions (Madhiraju et al., 2024).

- **Recency (R):** The number of days since the customer’s most recent purchase.
- **Frequency (F):** The total number of purchases made by the customer.
- **Monetary value (M):** The total amount of money spent by the customer.

After calculating these values, customers are scored or ranked. For example, a customer who purchased recently, buys frequently, and spends a large amount may receive high scores in all three categories (Christy et al., 2021).

5.2 Customer Groups in RFM

RFM analysis can be used to divide customers into groups such as best customers, loyal customers, new customers, at-risk customers, and lost customers. Table 1 shows examples of possible customer groups and suitable marketing actions.

Table 1: Examples of RFM-based customer groups

Customer Group	Typical RFM Pattern	Possible Marketing Action
Best customers	Recent purchase, high frequency, high spending	Loyalty rewards and premium offers
Loyal customers	Frequent purchases with moderate spending	Cross-selling and membership benefits
New customers	Recent purchase but low frequency	Welcome offers and onboarding messages
At-risk customers	Old purchase and low activity	Re-engagement emails or discounts
Lost customers	No recent purchase and low spending	Win-back campaigns

The main advantage of RFM analysis is that each customer group has a clear business meaning. Marketers can quickly identify customers who need rewards, attention, or reactivation campaigns (Christy et al., 2021; Alves Gomes and Meisen, 2023).

6 Clustering for Customer Segmentation

6.1 Concept of Clustering

Clustering is an unsupervised machine-learning method used to group similar objects together. In customer segmentation, each customer is represented using selected variables such as purchase frequency, average order value, product category preference, website visits, or RFM scores (John et al., 2023).

Unlike RFM analysis, clustering does not require predefined customer categories. Instead, it identifies natural groups present in the data. This makes clustering useful when businesses want to discover hidden patterns in customer behaviour.

6.2 K-Means Clustering

K-Means is a popular clustering algorithm. It divides customers into a chosen number of clusters, known as k . The algorithm works through the following steps:

1. Select the number of clusters, k .
2. Choose initial cluster centres.
3. Assign each customer to the nearest cluster centre.
4. Recalculate the centre of each cluster.
5. Repeat the process until cluster membership becomes stable.

The value of k is commonly selected using methods such as the Elbow Method or Silhouette Score. These methods help identify a suitable number of customer groups based on cluster separation and cohesion (John et al., 2023).

7 Comparative Analysis

Table 2 presents a comparison between RFM analysis and clustering.

RFM analysis is easier for marketing teams to understand because its variables directly represent customer purchasing behaviour (Christy et al., 2021). In contrast, clustering is more flexible because it can use several customer characteristics at the same time. However, clustering results may be difficult to explain if the selected variables are complex or poorly prepared (John et al., 2023).

Table 2: Comparison of RFM analysis and clustering

Aspect	RFM Analysis	Clustering
Main idea	Ranks customers using recency, frequency, and spending	Groups customers based on similarity in selected variables
Ease of use	Simple and easy to apply	Requires technical knowledge and data preparation
Interpretability	Very high because groups have clear meanings	Can be high, but cluster meaning must be analysed carefully
Data required	Customer ID, purchase date, purchase amount	Can use many behavioural and transactional variables
Flexibility	Limited to three main purchase measures	High, as many variables can be included
Hidden patterns	May miss complex customer behaviour	Can identify less obvious customer groups
Marketing use	Useful for loyalty, retention, and win-back campaigns	Useful for personalized and data-driven marketing
Main challenge	May oversimplify customer behaviour	Choosing suitable variables and number of clusters

8 Discussion

RFM analysis and clustering should not be viewed only as competing methods. In many practical cases, they are used together. A common approach is to calculate RFM values first and then apply K-Means clustering to those values. This approach retains the simplicity of RFM while using clustering to identify natural customer groups (Alves Gomes and Meisen, 2023; John et al., 2023).

For example, an online retail company may calculate RFM scores for all customers and then apply K-Means clustering to divide them into three groups: high-value active customers, occasional buyers, and inactive customers. The company can then design separate marketing campaigns for each group. This combined approach can provide more actionable customer segments than using RFM scores alone.

Clustering alone may be more suitable when a business has rich data beyond purchase history, such as browsing behaviour, product preferences, customer complaints, or app usage. In such cases, RFM may not fully represent customer behaviour because it mainly focuses on past purchases (Alves Gomes and Meisen, 2023; John et al., 2023).

9 Limitations

This paper is based on a review of existing literature rather than the analysis of a new customer dataset. Therefore, it does not test the performance of RFM and clustering on a specific business dataset.

RFM analysis has some limitations. It mainly considers past purchasing behaviour and may not include customer satisfaction, preferences, demographics, or changing behaviour over time. Clustering also has limitations. Its results depend strongly on data quality, variable selection, scaling, and the chosen number of clusters. Poor data preparation can lead to unclear or misleading customer groups (John et al., 2023).

10 Conclusion

RFM analysis and clustering are both valuable methods for customer segmentation. RFM analysis is simple, intuitive, and directly related to the customer's purchasing behavior. It is appropriate for companies that want to quickly segment their customers and/or develop marketing strategies for these segments, such as loyalty programs, retention tactics, and win-back strategies.

Meanwhile, the clustering method, particularly K-Means, is more flexible and can find unobserved patterns in the data. It is suitable for companies that want to do a thorough customer analysis and build detailed marketing strategies based on the observed patterns.

Depending on the company's needs, one method might be more appropriate than the others. However, in most cases, the analysis is more effective when using both methods together. K-Means can be used to group customers' RFM scores into segments. Overall, RFM analysis is a valuable tool for any company that wants to understand its customer base and develop a successful marketing strategy.

References

- Christy, A. J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM ranking—An effective approach to customer segmentation. *Journal of King Saud University—Computer and Information Sciences*, 33(10), 1251–1257. <https://doi.org/10.1016/j.jksuci.2018.09.004>
- Alves Gomes, M., & Meisen, T. (2023). A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. *Information Systems and e-Business Management*, 21, 527–570. <https://doi.org/10.1007/s10257-023-00640-4>
- John, J. M., Shobayo, O., & Ogunleye, B. (2023). An exploration of clustering algorithms for customer segmentation in the UK retail market. *Analytics*, 2(4), 809–823. <https://doi.org/10.3390/analytics2040042>
- Lewaaelhamd, I. (2024). Customer segmentation using machine learning model: An ap-

plication of RFM analysis. *Journal of Data Science and Intelligent Systems*, 2(1), 165–172. <https://doi.org/10.47852/bonviewJDSIS32021293>

Uthara, A. S., Sajikumar, S., & Jisha, R. C. (2024). Customer segmentation using RFM model and a new approach to encourage loyal customers. In *Intelligent Strategies for ICT, Lecture Notes in Networks and Systems* (Vol. 941). Springer. https://doi.org/10.1007/978-981-97-1260-1_10

Madhiraju, B., Reddy, S., & Sasikala, G. (2024). Customer segmentation using RFM analysis. *EPRA International Journal of Economic and Business Review*, 12(7), 15–22. <https://doi.org/10.36713/epra17685>