

Customer Segmentation Using Clustering Techniques in Retail: A Narrative Literature Review

S. Kavita¹, Dr. V. Kathiresan²

¹Research Scholar, Department of Computer Science, A.V.P. College of Arts and Science, Tiruppur, Tamil Nadu, India

²Associate Professor and Principal, A.V.P. College of Arts and Science, Tiruppur, Tamil Nadu, India

Corresponding author: S. Kavita, e-mail: kavitarajeshphd@gmail.com

ORCID: 0009-0002-4367-1376

Abstract

Purpose Customer segmentation is central to retail decision-making, and clustering-based approaches built on the RFM (Recency-Frequency-Monetary) framework are the dominant technique family used to derive it. This paper presents narrative literature review that synthesises research on clustering-based customer segmentation in retail, tracing the evolution of the RFM framework, by comparing clustering algorithms and validation methods, and surveying the deep emerging learning approaches.

Methods A narrative non-systematic review methodology is used. Relevant peer-reviewed articles, conference proceedings, and book chapters were identified through targeted searches across major themes: RFM model foundations and extensions, clustering algorithms, clustering validation metrics, segmentation based on deep learning, and sector-specific applications. More than fifty sources which were meeting these themes were reviewed and synthesised.

Findings The RFM framework used has been extended along many dimensions (LRFM, RFMT, RFMTC, RFM-D, LRFMV) to capture loyalty duration, churn risk, product diversity, and purchase volume and the K-Means remains the dominant clustering algorithm. The hybrid and density-based methods, along with emerging autoencoder-based deep clustering approaches, are increasingly used to handle non-linear and high-dimensional customer data. Validation practice is inconsistent across studies, like with the silhouette score, Davies-Bouldin index, and Calinski-Harabasz index each used but rarely reported together, limiting comparability across findings.

Conclusion In spite of three decades of active research, clustering-based customer segmentation in retail continues to evolve along largely separate tracks — feature engineering (RFM extensions), algorithm selection, and validation methodology — with limited standardisation across these studies. This review identifies this fragmentation as the central open gap of study and outlines directions far more integrated, rigorously validated segmentation research.

Keywords customer segmentation · clustering · RFM analysis · K-Means · retail analytics · literature review

1. Introduction

Customer segmentation — the fundamental practice of dividing a customer base into groups that share similar characteristics or behaviors — is a foundational technique in retail analytics, behind targeted marketing, inventory planning, and customer relationship management [1]. Among the various customer-segmentation approaches, clustering methods based on the Recency, Frequency, and Monetary (RFM) framework have emerged as one of the most widely adopted and influential paradigms in both academic research and business practice. Their popularity stems from their intuitive and easily interpretable structure, reliance on readily available transactional data, and proven ability to generate actionable customer groups that support targeted marketing, customer-retention strategies, personalized offerings, and improved business decision-making across diverse sectors. [2, 4].

Research in this area spans largely on distinct strands: the evolution of RFM into richer behavioral feature sets; the selection comparison of clustering algorithms; the validation of resulting clusters; and, more recently, the application of deep learning that helps to overcome the limitations of traditional methods on complex, high-dimensional retail data. While each of these strands is individually well developed, they are rarely synthesised together, making it difficult for researchers and practitioners to see how advances in one area relate to established practice in another.

This review addresses that gap by synthesising the literature on clustering-based customer segmentation in retail across four thematic areas, which are

- (i) The evolution of the RFM framework and its extensions;
- (ii) The clustering algorithms used to for segmentation;
- (iii) The methods used to validate resulting clusters. T
- (iv) The emerging role of deep learning in segmentation.

The review concludes by highlighting persistent gaps that remain unaddressed in this body of work and by outlining promising directions for future research.

The remainder of this paper is organized as follows. Section 2 traces the evolution of the RFM framework. Section 3 reviews clustering algorithms applied to customer segmentation. Section 4 examines cluster validation methods. Section 5 surveys deep learning-based segmentation approaches. Section 6 discusses sector-specific applications. Section 7 synthesises research gaps and future directions, and Section 8 concludes the review.

2. Evolution of the RFM Framework

The RFM model, which originated in direct-marketing practice and was widely popularized by Hughes, evaluates customers along three dimensions: **Recency** (time since the most recent purchase), **Frequency** (number of purchases within a specified period), and **Monetary value** (total or average customer expenditure). Its enduring appeal lies in its reliance on transactional data that most retailers routinely collect, enabling broad application without substantial additional data-collection costs.[4].

A substantial body of research has extended the basic RFM model to capture behavioral dimensions that it does not inherently represent. Reinartz and Kumar [5] observed that RFM alone cannot distinguish between long-term and short-term customer relationships, thereby motivating relationship-length-aware extensions. The LRFM model—comprising Length, Recency, Frequency, and Monetary value—addresses this limitation by incorporating the duration of the customer relationship. It has been applied to market segmentation in a children’s dental clinic [6] and to patient-value analysis in a pediatric dental-clinic case study. In these applications, self-organizing maps (SOMs), often combined with K-means clustering, were used to identify cluster structures before customer or patient profiles were interpreted. Further extensions have incorporated additional behavioral and economic dimensions into the RFM framework. Peker, Koçyiğit [7], and Eren proposed the LRFMP model—Length, Recency, Frequency, Monetary value, and Periodicity—for customer segmentation in the grocery retail industry. By adding Periodicity, defined as the interval or regularity of customers’ purchases, the model captures purchasing rhythms that are not reflected in conventional RFM or LRFM measures. The study used transaction data from a Turkish grocery chain and K-means clustering to identify five customer groups, including high-contribution loyal customers, uncertain customers, and high-spending lost customers. Yeh, Yang, and Ting [8] proposed RFMTC, adding Time-since-first-purchase and a Churn-probability estimate derived from a Bernoulli sequence model, enabling segmentation to directly incorporate retention risk. Soeini and Fathalizade extended RFM for the insurance industry by adding duration and cost dimensions [applied in 41], while Mahfuza et al. [9] proposed LRFMV, adding a volume dimension to capture the relationship between purchase quantity and profit in superstore data. Beyond behavioral extensions, several studies have combined RFM with other analytical techniques rather than adding new variables. Cheng and Chen [10] combined RFM scoring with rough set theory to derive classification rules for customer value segments. Christy et al. [2] proposed an RFM ranking approach as an alternative to clustering for deriving segment order directly from RFM scores. Guney, Peker, and Turhan [11] combined an LRFMP-based clustering stage with association rule mining in a video-on-demand context, illustrating a rare example of the segmentation-plus-pattern-mining integration that remains uncommon in the broader literature (see Section 7). Most recently, Ozcan [12] proposed an extended RFM model incorporating price sensitivity, shopping regularity, and basket depth for e-commerce segmentation, reporting five behaviorally distinct clusters when combined with K-Means.

Coussement, Van den Bossche, and De Bock [13] benchmarked RFM-based segmentation against logistic regression and decision trees, finding that data accuracy substantially affects segmentation performance regardless of technique—is a methodological caution that is echoed, but not always addressed, across the wider RFM literature reviewed here.

3. Clustering Algorithms for Customer Segmentation

3.1 Partitioning and Density-Based Methods

K-Means, introduced by MacQueen [14], remains by a huge margin the most frequently applied algorithm for RFM-based customer segmentation, owing to its computational efficiency and ease of interpretation [15, 20]. Its main limitation is its sensitivity to the initial choice of k , sensitivity to outliers, and an assumption of roughly spherical, similarly sized clusters have motivated a range of alternative and hybrid approaches.

K-Medoids addresses K-Means' sensitivity to outliers' anomaly by using actual data points (medoids) rather than centroids as cluster representatives, at higher computational cost [16]. Wu, Chang, and Lo [15] applied both RFM-based K-Means and related partitioning methods to customer value analysis for a retail outfitter. Mini-Batch K-Means, evaluated alongside standard K-Means and K-Medoids by Mahfuza et al. [9], trades a small amount of clustering quality for substantially reduced computation time on large transactional datasets, making it attractive for retailers with high transaction volumes.

Density-based methods, principally DBSCAN [17], cluster points which is based on local density rather than distance to a centroid, allowing them to detect arbitrarily shaped clusters and explicitly identify outliers rather than forcing every customer into a segment. Monalisa [18] applied DBSCAN specifically to outlier detection within RFM and LRFM feature spaces, arguing that outlier customers (for example, occasional very-high-value purchasers) are frequently mis-segmented by centroid-based methods and are better handled by explicit outlier identification prior to clustering.

3.2 Fuzzy and Hierarchical Methods

Where strict cluster boundaries are unrealistic for instance, a customer near the boundary between “loyal” and “at-risk” segments — fuzzy clustering methods assign partial membership across multiple clusters. Zahrotun [19] applied Fuzzy C-Means to RFM-derived customer data in an online shop context, reporting that fuzzy membership scores provided more nuanced customer profiles than the hard assignments produced by K-Means. This approach is particularly useful in customer segmentation because purchasing behaviour often overlaps across customer groups. Hierarchical (agglomerative) clustering has also been used, either as a standalone method or as a cross-validation check against partitioning results, since it does not require the number of clusters to be specified in advance and produces a dendrogram that can itself inform the choice of k [20, 21].

3.3 Algorithm Comparison Studies

A recurring theme across the literature is direct comparison of multiple algorithms on the same dataset rather than commitment to a single method. Arun [22] compared Hierarchical Clustering, K-Means, DBSCAN, and Fuzzy C-Means on evaluation metrics including silhouette score and Davies-Bouldin index, finding that no single algorithm dominated across all metrics — hierarchical clustering produced the best-separated clusters by silhouette score, while a different algorithm minimized the Davies-Bouldin index, underscoring that algorithm choice is metric- and context-dependent rather than universal. This pattern of context-dependent algorithm performance recurs across the comparison studies reviewed here and is discussed further in Section 4.

Corrected till here

4. Cluster Validation Methods

Because clustering is unsupervised, there is no ground-truth label against which to directly measure accuracy; instead, studies rely on internal validation metrics that assess cluster cohesion and separation using the data itself. Three metrics dominate the reviewed literature: the silhouette coefficient [23], the Davies-Bouldin index [24], and the Calinski-Harabasz index [25].

The silhouette coefficient, for each point, compares its average distance to other points in its own cluster against its average distance to points in the nearest neighboring cluster, producing a score between -1 and

+1 where higher values indicate better-defined clusters [23]. It is the most widely reported metric across the studies reviewed here [9, 22], in part because its bounded range gives it an intuitive interpretation that the Davies-Bouldin and Calinski-Harabasz indices lack.

The Davies-Bouldin index measures the average similarity between each cluster and its most similar counterpart, with lower values indicating better separation [24]; unlike the silhouette score it has no fixed upper bound, complicating cross-study comparison. The Calinski-Harabasz index, based on the ratio of between-cluster to within-cluster dispersion, is used less frequently in the retail segmentation literature reviewed here but appears alongside the other two in more recent comparative studies [22].

A persistent limitation across the reviewed literature is that few studies report more than one or two of these metrics together, and even fewer justify why a given metric was chosen over the alternatives. Arun's [22] finding that different metrics favored different algorithms on the same dataset illustrates the practical consequence of this inconsistency: conclusions about “which algorithm is best” in the literature are often metric-dependent rather than dataset-dependent, a point rarely made explicit in the individual studies themselves. The elbow method (based on within-cluster sum of squares) remains widely used specifically for selecting the number of clusters k , typically prior to — and independently of — the internal validation metrics used to assess the resulting clusters [9, 26].

5. Deep Learning-Based Segmentation

Traditional clustering algorithms assume that meaningful structure exists in the original feature space, an assumption that becomes increasingly strained as retailers incorporate higher-dimensional behavioral data — browsing sequences, review text, product-level purchase histories — alongside or instead of simple RFM scores. This has motivated a fast-growing body of work applying deep learning to customer segmentation, concentrated almost entirely in the last two to three years of the literature reviewed here.

Autoencoders — neural networks trained to reconstruct their own input through a compressed intermediate representation — are the most commonly applied deep learning technique in this space, used primarily for non-linear dimensionality reduction prior to conventional clustering [27, 28]. Chavan et al. [28] combined autoencoder-based dimensionality reduction with conventional clustering on a large retail dataset, reporting that the hybrid approach uncovered customer groupings not visible when clustering was applied directly to the original high-dimensional feature space. Madanchian, Farhaoui, and Taherdoost [29] reviewed autoencoders and generative models for segmentation across retail and financial-services case studies, concluding that these methods improve segmentation scalability and precision but require substantially more data and computational resources than RFM-based clustering.

A related and more tightly integrated approach is deep embedded clustering, in which the dimensionality-reduction and clustering objectives are optimized jointly rather than sequentially. Xie, Girshick, and Farhadi [30] introduced this approach in the general machine learning literature, and Yang et al. [31] extended it with a “K-Means-friendly” embedding space that explicitly shapes the learned representation to suit centroid-based clustering. Li et al. [32] proposed a residual-autoencoder-based deep embedded clustering variant, reporting improved cluster compactness over standard autoencoder-plus-K-Means pipelines. Berahmand et al. [27] surveyed autoencoder applications across machine learning broadly, providing a methodological anchor for the retail-specific applications discussed here.

Despite this growing body of work, deep learning-based segmentation studies in the retail literature reviewed here rarely report the same internal validation metrics (silhouette score, Davies-Bouldin index) used in the traditional-clustering literature discussed in Section 4, and even more rarely compare directly against an RFM-plus-K-Means baseline on the same dataset. This makes it difficult to assess, from the literature alone, how much of the reported improvement is attributable to the deep learning component specifically rather than to dataset or evaluation differences across studies — a gap discussed further in Section 7.

6. Sector-Specific Applications

While retail is the primary focus of this review, RFM-based clustering has been applied across a wide range of adjacent sectors, and this cross-sector literature has directly shaped methodological practice within retail-specific studies. In banking, Hsieh [33] combined data mining with behavioral scoring for customer analysis, while Sohrabi and Khanlari [34] used RFM to measure customer lifetime value in a financial-services context. In telecommunications, Li, Shue, and Lee [35] applied RFM-informed business intelligence to strategy-making for internet service providers.

Healthcare applications include Mohammadzadeh, Hoseini, and Derafshi's [36] use of RFM to model patient churn behavior in a specialized clinic setting, and the dental-clinic LRFM case studies discussed in Section 2 [6, 38]. Hospitality and travel applications include Lumsden, Beldona, and Morrison's [37] application of the RFM framework to customer value in an all-inclusive travel vacation club. Even political campaign management has adopted the framework: Nagesh [41] used RFM scoring combined with XGBoost-based cluster prediction to generate political engagement scores.

This breadth of application underscores a consistent finding across the reviewed literature: the core RFM-plus-clustering pattern generalizes well across domains with transactional or engagement data, even as the specific extensions and clustering algorithms used vary by sector. Retail-specific studies (Sections 2 and 3) tend to emphasize purchase volume and product diversity extensions, whereas the sector applications reviewed here more often emphasize churn and lifetime-value extensions — a pattern of domain-specific specialization that itself represents an implicit, largely unstated research gap: retail segmentation studies rarely draw on churn-modeling techniques well established in banking and telecom, and vice versa.

7. Research Gaps and Future Directions

Synthesising the literature reviewed in Sections 2 through 6 surfaces three related gaps that cut across the individual families of technique.

Fragmentation across feature engineering, algorithm choice, and validation. Studies reviewed here classically innovate along exactly one of these three axes — a new RFM extension applied with standard K-Means and a single validation metric [8, 9, 12]; a new algorithm applied to standard RFM features [18, 19, 22]; or a new validation approach applied post hoc to an existing pipeline. Few studies vary more than one axis at a time, making it difficult to isolate which component — features the algorithm, or the validation choice — actually drives reported improvements.

Inconsistent and incomplete validation reporting. As discussed in Section 4, most previous studies report only one or two internal validation metrics, and Arun's [22] finding that the algorithm rankings

differ by metric, actually suggests that single-metric conclusions throughout this literature should be treated cautiously. No study reviewed here reports external validation (for example, business-outcome-based validation such as measured campaign lift per segment) alongside internal metrics, leaving the practical business value of statistically “better” clusters largely unverified.

Limited integration with downstream pattern discovery. With the exception of Guney, Peker, and Turhan's [11] combined clustering-and-association-rule-mining approach in a video-on-demand setting, the segmentation literature reviewed here treats customer clustering as an endpoint rather than as an input to further analysis such as product-level association rule mining. This is consistent with the broader disconnect noted in prior retail data-mining literature between customer-level segmentation and product-level pattern discovery techniques, which are more often studied independently than in combination. These gaps motivate research that deliberately combines a validated feature-engineering choice, a justified (not merely convenient) algorithm selection, multi-metric validation reported transparently, and — where relevant — explicit integration with downstream analytical tasks such as association rule mining, rather than treating segmentation as an isolated technique applied and reported in a methodological vacuum.

8. Conclusion

This review synthesised the literature on clustering-based customer segmentation in retail across four themes: the evolution of the RFM framework and its many extensions, the clustering algorithms used to systematize segmentation, the methods used to validate resulting clusters, and the emerging role of deep learning. More than three decades after Hughes [3] first formalized RFM, the framework continues to be actively extended, most recently toward e-commerce-with specific behavioral dimensions [12] and deep-learning-based on feature representations [27–32].

The review's central finding is one of fragmentation rather than absence of progress: individual technique families — RFM extensions, algorithm comparisons, validation methodology, and deep learning — are each reasonably mature, but are rarely combined or evaluated against one another within a single, methodologically transparent study. Addressing this fragmentation, through studies that combine validated feature engineering, justified algorithm selection, multi-metric validation, where appropriate, integration with downstream analytical tasks, represents the most direct path toward more reliable and more actionable customer segmentation research in retail.

Declarations

Funding The authors did not receive support from any organization for the submitted work. [Authors: please edit this statement if funding was received.]

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethics Approval Not applicable. This is a literature review and did not involve human participants, human data, or biological material.

Data Availability Not applicable. No new datasets were generated or analysed; all sources reviewed are cited in the reference list.

Author Contributions S. Kavita: conceptualization, literature search, analysis, manuscript drafting. Dr. V. Kathiresan: supervision, review and editing. Both authors read and approved the final manuscript. [Authors: please edit to reflect actual contributions.]

References

1. Desouza KC, Awazu Y, Jha S, Dombrowski C, Papagari S, Baloh P, et al. Customer-driven innovation. *Res Technol Manag.* 2008;51(3):35–44.
2. Christy AJ, Umamakeswari A, Priyatharsini L, Neyaa A. RFM ranking: an effective approach to customer segmentation. *J King Saud Univ Comput Inf Sci.* 2021. <https://doi.org/10.1016/j.jksuci.2018.09.004>
3. Hughes AM. Database marketing. Chicago: Probus Publishing Company; 1994.
4. Peker S, Kocyigit A, Eren PE. LRFMP model for customer segmentation in the grocery retail industry: a case study. *Mark Intell Plan.* 2017;35(4):544–59. <https://doi.org/10.1108/MIP-11-2016-0210>
5. Reinartz WJ, Kumar V. On the profitability of long-life customers in a noncontractual setting: an empirical investigation and implications for marketing. *J Mark.* 2000;64(4):17–35. <https://doi.org/10.1509/jmkg.64.4.17.18077>
6. Wei JT, Lin SY, Weng CC, Wu HH. A case study of applying LRFM model in market segmentation of a children's dental clinic. *Expert Syst Appl.* 2012;39(5):5529–33. <https://doi.org/10.1016/j.eswa.2011.11.066>
7. Sarvari P, Ustundag A, Takci H. Performance evaluation of different customer segmentation approaches based on RFM and demographics analysis. *Kybernetes.* 2016;45(7):1129–57. <https://doi.org/10.1108/K-07-2015-0180>
8. Yeh IC, Yang KJ, Ting TM. Knowledge discovery on RFM model using Bernoulli sequence. *Expert Syst Appl.* 2009;36(3):5866–71. <https://doi.org/10.1016/j.eswa.2008.07.018>
9. Mahfuza R, Islam N, Toyeb M, Emon MAF, Chowdhury SA, Alam MGR. LRFMV: an efficient customer segmentation model for superstores. *PLOS ONE.* 2022;17(12):e0279262. <https://doi.org/10.1371/journal.pone.0279262>
10. Cheng CH, Chen YS. Classifying the segmentation of customer value via RFM model and RS theory. *Expert Syst Appl.* 2009;36(3):4176–84. <https://doi.org/10.1016/j.eswa.2008.04.003>
11. Guney S, Peker S, Turhan C. A combined approach for customer profiling in video on demand services using clustering and association rule mining. *IEEE Access.* 2020;8:84326–35.
12. Ozcan T. Customer segmentation using an extended RFM model and clustering algorithms in e-commerce. *J Theor Appl Electron Commer Res.* 2026;21(5):142. <https://doi.org/10.3390/jtaer21050142>
13. Coussement K, Van den Bossche FAM, De Bock KW. Data accuracy's impact on segmentation performance: benchmarking RFM analysis, logistic regression, and decision trees. *J Bus Res.* 2014;67(12):2751–8. <https://doi.org/10.1016/j.jbusres.2012.09.024>
14. MacQueen J. Some methods for classification and analysis of multivariate observations. In: *Proc 5th Berkeley Symp Math Stat Probab.* 1967;1:281–97.
15. Wu HH, Chang EC, Lo CF. Applying RFM model and K-means method in customer value analysis of an outfitter. In: *Proc 16th ISPE Int Conf Concurrent Eng.* p. 665–72.

16. Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. New York: Wiley; 1990.
17. Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proc 2nd Int Conf Knowledge Discovery Data Mining (KDD). 1996.
18. Monalisa S. Analysis outlier data on RFM and LRFM models to determining customer loyalty with DBSCAN algorithm. In: Proc Int Symp Adv Intell Inform (SAIN). 2018. p. 1–5.
<https://doi.org/10.1109/SAIN.2018.8673380>
19. Zahrotun L. Implementation of data mining technique for customer relationship management (CRM) on online shop tokodipers.com with fuzzy c-means clustering. In: Proc 2017 2nd Int Conf Inf Technol Inf Syst Electr Eng (ICITISEE). Yogyakarta; 2017. p. 299–303.
20. Bachtiar FA. Customer segmentation using two-step mining method based on RFM model. In: Proc Int Conf Sustain Inf Eng Technol (SIET). p. 10–5.
21. Tavakoli M, Molavi M, Masoumi V, Mobini M, Etemad S, Rahmani R. Customer segmentation and strategy development based on user behavior analysis, RFM model and data mining techniques: a case study. In: Proc 2018 IEEE 15th Int Conf E-Business Eng (ICEBE). 2018.
22. Arun AP. Optimizing customer segmentation: a comparative analysis of clustering algorithms using evaluation metrics. 2024.
23. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53–65.
24. Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intell. 1979;1(2):224–7.
25. Calinski T, Harabasz J. A dendrite method for cluster analysis. Commun Stat. 1974;3(1):1–27.
26. Li J, Song S, Zhang Y, Zhou Z. Robust K-median and K-means clustering algorithms for incomplete data. Math Probl Eng. 2016;ID 4321928. <https://doi.org/10.1155/2016/4321928>
27. Berahmand K, Daneshfar F, Salehi ES, Li Y, Xu Y. Autoencoders and their applications in machine learning: a survey. Artif Intell Rev. 2024;57(2):28.
28. Chavan A, Rathod I, Guo B, Ren R, Kang T, Jung Y. Deep learning strategies for optimizing customer segmentation. In: AMCIS 2025 Proceedings. 2025.
29. Madanchian M, Farhaoui Y, Taherdoost H. Enhancing customer segmentation with unsupervised deep learning. In: Proc ICAISE 2024. Cham: Springer; 2025.
30. Xie J, Girshick R, Farhadi A. Unsupervised deep embedding for clustering analysis. In: Proc 33rd Int Conf Mach Learn (ICML). 2016.
31. Yang B, Fu X, Sidiropoulos ND, Hong M. Towards k-means-friendly spaces: simultaneous deep learning and clustering. In: Proc 34th Int Conf Mach Learn (ICML). 2017.
32. Li M, Cao C, Li C, Yang S. Deep embedding clustering based on residual autoencoder. Neural Process Lett. 2024;56(2):127.
33. Hsieh NC. An integrated data mining and behavioral scoring model for analyzing bank customers. Expert Syst Appl. 2004;27(4):623–33. <https://doi.org/10.1016/j.eswa.2004.06.007>
34. Sohrabi B, Khanlari A. Customer lifetime value (CLV) measurement based on RFM model. Iran Account Audit Rev. 2007;14(47):7–20.
35. Li ST, Shue LY, Lee SF. Business intelligence approach to supporting strategy-making of ISP service management. Expert Syst Appl. 2008;35(3):739–54.
<https://doi.org/10.1016/j.eswa.2007.07.049>

36. Mohammadzadeh M, Hoseini ZZ, Derafshi H. A data mining approach for modeling churn behavior via RFM model in specialized clinics: case study of a public sector hospital in Tehran. *Procedia Comput Sci.* 2017;120:23–30. <https://doi.org/10.1016/j.procs.2017.11.206>
37. Lumsden SA, Beldona S, Morrison AM. Customer value in an all-inclusive travel vacation club: an application of the RFM framework. *J Hosp Leis Mark.* 2008;16(3):270–85. <https://doi.org/10.1080/10507050801946858>
38. Wu HH, Lin SY, Liu CW. Analyzing patients' values by applying cluster analysis and LRFM model in a pediatric dental clinic in Taiwan. *ScientificWorldJournal.* 2014;2014:685495. <https://doi.org/10.1155/2014/685495>
39. Soeini RA, Fathalizade E. Customer segmentation based on modified RFM model in the insurance industry.
40. Etzion O, Fisher A, Wasserkug S. e-CLV: a modeling approach for customer lifetime evaluation in e-commerce domains, with an application and case study for online auction. *Inf Syst Front.* 2005;7:421–34. <https://doi.org/10.1007/s10796-005-4812-6>
41. Nagesh JM. Generating political scores using RFM model and cluster prediction by XGBoost. In: *Proc 2019 Int Conf Comput Sci Comput Intell (CSCI).* Las Vegas; 2019. p. 1333–6.
42. [Author names to be verified against source]. How can algorithms help in segmenting users and customers? A systematic review and research agenda for algorithmic customer segmentation. *J Mark Anal.* 2023;11:677–92. <https://doi.org/10.1057/s41270-023-00235-5>
43. [Author names to be verified against source]. An exploration of clustering algorithms for customer segmentation in the UK retail market. *Analytics.* 2023;2(4):809–23.
44. Ho T, Nguyen S, Nguyen H, Nguyen N, Man DS, Le TG. An extended RFM model for customer behaviour and demographic analysis in retail industry. 2023.
45. Shirole R, Salokhe L, Jadhav S. Customer segmentation using RFM model and K-means clustering. *Int J Sci Res Sci Technol.* 2021;8:591–7.
46. Talaat FM, Aljadani A, Alharthi B, Farsi MA, Badawy M, Elhosseini M. A mathematical model for customer segmentation leveraging deep learning, explainable AI, and RFM analysis in targeted marketing. 2023.
47. Sarkar M, Puja AR, Chowdhury FR. Optimizing marketing strategies with RFM method and K-means clustering-based AI customer segmentation analysis. 2024.
48. Kasem MS, Hamada M, Taj-Eddin I. Customer profiling, segmentation, and sales prediction using AI in direct marketing. 2024.
49. Joung J, Kim H. Interpretable machine learning-based approach for customer segmentation for new product development from online product reviews. 2023.
50. Wang CH. Outlier identification and market segmentation using kernel-based clustering techniques. *Expert Syst Appl.* 2009;36(2):3744–50. <https://doi.org/10.1016/j.eswa.2008.02.037>
51. Pakyurek M, Sezgin MS, Kestepi S, Bora B, Duzagac R, Yildiz OT. Customer clustering using RFM analysis. In: *Proc 26th Signal Process Commun Appl Conf (SIU).* 2018. p. 1–4.