Transactions on Soft Computing

www.tsc.reapress.com

Trans. Soft. Comput. Vol. 1, No. 1 (2025) 14–26.

Paper Type: Original Article

A Comprehensive RFM-A Framework: Integrating Age for Enhanced Customer Segmentation and Marketing Strategies

Mansoureh Naderipour[1,*], Hamed Hosseini Nasabian[1]

[1] Department of Industrial Engineering, Faculty of Engineering, Shahid Bahonar University of Kerman, Kerman, Iran; M.naderipour@uk.ac.ir; HamedHosseini@eng.uk.ac.ir.

Citation:

Received: 10 September 2024 *Revised: 14 October 2024* *Accepted: 05 January 2025*	Naderipour, M., Hosseini Nasabian, H. (2025). A comprehensive RFM-A framework: Integrating age for enhanced customer segmentation and marketing strategies. *Transactions on soft computing, 1*(1), 14-26.

Abstract

Customer segmentation is seen as one of the pillars of a successful advertising campaign. Marketers give great importance to this flagship phase in the process of new products marketing. Successful segmentation will involve successful “Customer Targeting” and therefore a profitable customer marketing campaign. Many works have dealt with customer segmentation by applying the famous Recency, Frequency and Monetary model. This model suffers from insufficiency by ignoring other important parameters according to the field of application. In this article, a new classification model is presented by adding the age ("A") as the fourth parameter, referring to the age of customers. The segmentation based on RFM-A is applied in a retail market in order to detect behavior patterns for a customer. The proposed model increases the quality of the prediction of customer behavior and Companies could predict, customers who will respond positively.

Keywords: RFM, Segmentation, Customer, Age, Marketing, Prediction.

1 | Introduction

In the dynamic landscape of modern commerce, businesses are increasingly turning to data-driven strategies to enhance their competitive edge, optimize operational efficiency, and foster customer loyalty [1], [2]. The advent of big data and advanced analytics has revolutionized marketing practices, enabling organizations to extract actionable insights from vast datasets. Among these, customer segmentation stands out as a pivotal technique, allowing companies to categorize their customer base into distinct groups based on behavioral and demographic characteristics. This segmentation facilitates the design of targeted marketing campaigns,

Corresponding Author: M.naderipour@uk.ac.ir
https://doi.org/10.48314/tsc.v1i1.34

personalized product offerings, and improved customer experiences, ultimately driving revenue growth and customer retention [3], [4].

The Recency, Frequency, Monetary (RFM) model has long been a cornerstone of customer segmentation, providing a robust framework to evaluate customers based on their purchasing behavior [5]. Recency measures the time elapsed since a customer's last purchase, Frequency counts the number of purchases within a specific period, and Monetary value quantifies the total amount spent. By analyzing these metrics, businesses can identify high-value customers, predict purchasing patterns, and tailor marketing strategies accordingly [6]. For instance, RFM analysis has been successfully applied in retail to distinguish between loyal customers who frequently purchase high-value items and occasional buyers who contribute less to revenue [7]. However, despite its widespread adoption, the traditional RFM model has notable limitations, particularly its lack of consideration for demographic factors that significantly influence consumer behavior.

One critical demographic variable often overlooked by the RFM model is customer age, which plays a pivotal role in shaping purchasing preferences, brand engagement, and response to marketing campaigns [8]. For example, younger customers (Aged 15–34) may prioritize trendy, technology-driven products and rely heavily on social media recommendations, while older customers (Aged 55–67) may value cost-effective, health-oriented products and prefer in-store experiences [9]. Ignoring age in segmentation can lead to overly generalized marketing strategies that fail to resonate with specific customer groups, reducing campaign effectiveness and customer satisfaction. To address this gap, this study proposes the RFM-A model, an innovative extension of the RFM framework that incorporates customer age as a fourth dimension. By integrating age, the RFM-A model enables businesses to achieve more granular segmentation, uncover nuanced behavioral patterns, and design marketing campaigns that align with the preferences of different age groups.

The importance of incorporating demographic variables such as age into customer segmentation cannot be overstated, particularly in the context of e-commerce, where diverse customer bases interact with online platforms daily [2]. For instance, a retail company targeting tech-savvy millennials may benefit from offering personalized promotions on smartphones and gaming consoles. In contrast, a company catering to seniors may focus on healthcare products and loyalty discounts. The RFM-A model facilitates such tailored strategies by combining behavioral data (RFM) with demographic insights (Age), providing a holistic view of customer profiles. This approach not only enhances prediction accuracy but also enables businesses to optimize resource allocation, improve customer retention, and increase profitability [4].

The application of data science techniques, such as clustering and predictive modeling, further amplifies the effectiveness of the RFM-A model. By leveraging algorithms like K-Means clustering, businesses can group customers into homogeneous segments based on their RFM-A scores, enabling precise targeting of marketing efforts [6]. Moreover, the integration of age into the RFM framework aligns with recent trends in data-driven marketing, where machine learning and artificial intelligence are used to uncover complex patterns in consumer behavior [3]. For example, a study by Saura et al. [2] demonstrated that combining demographic and behavioral data improves the accuracy of demand forecasting and price optimization in Small and Medium-Sized Enterprises (SMEs).

This study applies the RFM-A model to a real-world dataset from an online retail store, comprising transaction records and customer demographics collected over one year (January 2024 to December 2024). The dataset includes approximately 10,000 transactions from 2,500 unique customers, with ages ranging from 15 to 67 years. By analyzing this data, this article aims to demonstrate the RFM-A model's ability to identify distinct customer segments, predict purchasing behavior, and inform strategic marketing decisions. The contributions of this study are threefold:

I. Introduction of the RFM-A model: A novel extension of the RFM framework is proposed by incorporating age as a critical variable, enhancing the granularity of customer segmentation.

II. Empirical validation: The RFM-A model is applied to a retail dataset using Python-based clustering techniques, validating its effectiveness through statistical analysis and visualizations.

III. Practical applications: Actionable recommendations for online retailers are provided to design age-specific marketing campaigns, improve customer retention, and optimize product offerings.

The remainder of this paper is organized as follows: Section 2 reviews related work on customer segmentation and RFM-based models. Section 3 details the methodology, including data preprocessing, RFM-A model implementation, and clustering techniques. Section 4 presents experimental results, including cluster analysis and visualizations, and discusses practical applications and limitations. And Section 6 concludes the study.

2 | Related Works

Customer segmentation has emerged as a critical strategy in modern marketing, enabling businesses to understand consumer behavior, tailor marketing campaigns, and optimize resource allocation [2]. The RFM model, introduced by Hughes [5], has been a cornerstone of customer segmentation, providing a straightforward yet effective framework to categorize customers based on their purchasing patterns. This section reviews the evolution of RFM-based models, their applications in various industries, and the growing emphasis on incorporating demographic variables, such as age, to enhance segmentation accuracy. The discussion is organized into three subsections: 1) traditional RFM models, 2) extensions of RFM with clustering techniques, and 3) integration of demographic variables in segmentation.

2.1 | Traditional RFM Models

The RFM model evaluates customers based on three key metrics: Recency (Time since the last purchase), Frequency (Number of purchases within a period), and monetary (Total spending) [5]. This model has been widely adopted in retail, e-commerce, and banking due to its simplicity and ability to identify high-value customers [6]. For instance, Saleem et al. [3] applied RFM analysis to e-commerce platforms, demonstrating its effectiveness in improving sales performance by identifying customers likely to respond to targeted promotions. Similarly, Siahaan and Sianipa [7] conducted a case study on retail store transaction data, using RFM to segment customers into groups such as loyal, occasional, and at-risk customers. Their findings highlighted the model's ability to predict purchasing behavior and inform marketing strategies, such as personalized discounts for frequent buyers.

Despite its strengths, the traditional RFM model has limitations, particularly its reliance on behavioral data alone, which may overlook critical demographic factors influencing consumer preferences [8]. For example, a young customer with high Frequency but low monetary value may exhibit different motivations than an older customer with similar RFM scores. This limitation underscores the need for enhanced models that incorporate additional variables to capture the complexity of consumer behavior.

2.2 | Extensions of RFM with Clustering Techniques

To address the limitations of traditional RFM models, researchers have integrated advanced data science techniques, such as clustering, to improve segmentation accuracy. Shirole et al. [6] combined RFM with K-Means clustering to segment customers in a retail setting, creating clusters such as "super customers," "Intermediate customers," and "At-risk customers" based on RFM scores. Their study demonstrated that clustering enhances the granularity of segmentation, enabling businesses to tailor marketing strategies to specific customer groups. Similarly, Ong [10] applied RFM analysis with K-Means clustering on a Snowpark platform, emphasizing the scalability of RFM-based models for large datasets. Their approach allowed retailers to process vast transaction records efficiently, identifying behavioral patterns that informed demand forecasting and inventory management.

Chandana [9] further explored the use of RFM analysis with K-Means clustering, advocating for the use of scoring systems to simplify segmentation. They argued that assigning scores (e.g., 1–5) to RFM variables, rather than using raw values, improves clustering outcomes by reducing the impact of outliers and normalizing data distributions. This scoring approach was efficient in identifying homogeneous customer groups, enabling businesses to design targeted campaigns with minimal computational complexity. However, these studies primarily focused on behavioral data, with limited attention to demographic variables that could further refine segmentation results.

2.3 | Integration of Demographic Variables in Segmentation

Recent research has highlighted the importance of incorporating demographic variables, such as age, gender, and location, into customer segmentation to capture nuanced behavioral patterns [2], [8]. Chugh [8] emphasized that customer behavior varies significantly across demographic groups, suggesting that RFM analysis could be enhanced by first segmenting customers based on demographics and then applying RFM for behavioral analysis. For example, a retail chain might segment customers by geographic region before applying RFM to identify high-value customers within each region. This two-stage approach improves the relevance of marketing strategies by aligning them with demographic-specific preferences.

Similarly, Chandana [9] proposed that integrating demographic information, such as age, into RFM analysis could enhance segmentation accuracy. They noted that age influences purchasing preferences, with younger customers often seeking trendy products and older customers prioritizing value and convenience. This insight is particularly relevant in e-commerce, where diverse customer bases interact with online platforms [4]. For instance, Pantano et al. [4] investigated the role of retail technology in reducing barriers for older consumers, finding that age-specific preferences significantly impact shopping behavior. Their study suggested that tailoring product offerings and user interfaces to older customers can improve engagement and loyalty.

Other studies have explored the integration of demographic and behavioral data in related contexts. Azad et al. [11] developed predictive models combining RFM metrics with social media interactions and demographic data, demonstrating improved accuracy in forecasting consumer purchase behavior. Likewise, Rane et al. [12] explored the use of demographic variables in metaverse-based marketing, highlighting the potential of age-informed strategies to enhance customer loyalty. These findings underscore the need for models like RFM-A, which explicitly incorporate age to address the limitations of traditional RFM frameworks.

2.4 | Relevance to the RFM-A Model

The proposed RFM-A model builds on the insights from these studies by integrating customer age as a fourth dimension in the RFM framework. Unlike traditional RFM models, which focus solely on behavioral metrics, RFM-A captures the interplay between age and purchasing behavior, enabling more precise segmentation [6], [7]. By combining age with recency, Frequency, and monetary metrics, the RFM-A model addresses the gap identified by Chugh [8] and Chandana [9], who emphasized the importance of demographic variables in segmentation. Furthermore, the use of clustering techniques, as demonstrated by Shirole et al. [6] and Ong [10], provides a robust methodology for implementing RFM-A, ensuring scalability and applicability to large datasets.

The RFM-A model also aligns with recent trends in data-driven marketing, where advanced analytics and machine learning are used to uncover complex consumer patterns [2], [3]. By incorporating age, the model enables businesses to design marketing campaigns that resonate with specific age groups, such as tech-savvy millennials or value-conscious seniors. This approach not only enhances customer engagement but also optimizes resource allocation, as businesses can focus on high-potential segments identified through RFM-A analysis.

In summary, the literature highlights the effectiveness of RFM-based models in customer segmentation, while also pointing to the need for integrating demographic variables to improve accuracy. The RFM-A model proposed in this study addresses this need by extending the RFM framework with age, offering a novel and practical solution for online retailers seeking to enhance their marketing strategies.

3 | Methodology

The RFM-A model extends the traditional RFM framework by incorporating customer age as a fourth dimension, enabling a more nuanced approach to customer segmentation in online retail. This section details the methodology employed to develop and validate the RFM-A model, including the dataset, data preprocessing steps, variable selection, model implementation, clustering techniques, and evaluation metrics. The methodology leverages Python-based tools and clustering algorithms to analyze customer transaction data, ensuring robust and reproducible results. The discussion is organized into five subsections: dataset description, data preprocessing, variable selection, RFM-A model implementation, and clustering and evaluation.

3.1 | Dataset Description

The dataset used in this study comprises transaction records from an online retail store, collected over a one-year period from January 2024 to December 2024. The dataset includes approximately 10,000 transactions from 2,500 unique customers, providing a rich source of behavioral and demographic data. Key attributes include customer ID, transaction date, purchase amount, and customer age, with ages ranging from 15 to 67 years. This diverse age range allows for a comprehensive analysis of purchasing behavior across different demographic groups, which is critical for validating the RFM-A model [2].

The dataset was sourced from the retailer's internal database, capturing transactions across various product categories, such as electronics, clothing, and healthcare products. Each transaction record includes metadata, such as product category and purchase channel (e.g., website or mobile app), which was used to contextualize customer behavior. The dataset's size and diversity make it suitable for clustering analysis, as it encompasses a wide range of purchasing patterns and demographic profiles [3]. To ensure data integrity, the dataset was verified for completeness and consistency before analysis, with any incomplete records flagged for preprocessing.

3.2 | Data Preprocessing

Data preprocessing is a critical step in ensuring the quality and reliability of the analysis, particularly when dealing with large and heterogeneous datasets [6]. The following preprocessing steps were applied to prepare the dataset for RFM-A analysis:

Handling missing values: Approximately 1.8% of the records had missing values for transaction dates or purchase amounts. These records were removed to avoid bias in the analysis, as imputation could introduce inaccuracies in time-sensitive metrics like recency [9].

Outlier detection and treatment: Purchase amounts were examined for outliers using the interquartile range (IQR) method. Values exceeding three standard deviations from the mean were capped at the 95th percentile to mitigate the impact of extreme purchases, such as bulk orders, which could skew monetary scores [7].

Normalization: To ensure equal weighting of RFM-A variables during clustering, all variables (RFM and age) were normalized to a 0–1 scale using min-max normalization. This step was essential to prevent variables with larger ranges (e.g., monetary) from dominating the clustering process [6].

Data aggregation: Transaction records were aggregated by customer ID to compute RFM-A metrics. For example, recency was calculated as the number of days since the customer's last purchase, and Frequency was determined by counting the number of transactions within one year.

These preprocessing steps ensured a clean and standardized dataset, suitable for accurate segmentation and clustering.

3.3 | Variable Selection

The RFM-A model is based on four key variables: RFM and Age. The selection of these variables was guided by their relevance to customer segmentation and their ability to capture both behavioral and demographic aspects of consumer behavior [8]. The rationale for each variable is as follows:

Recency (R): Measures the number of days since the customer's last purchase. Recent customers are more likely to respond to marketing campaigns, making recency a critical predictor of engagement [5].

Frequency (F): Counts the number of purchases within the analysis period. Frequent customers are typically more loyal and contribute significantly to revenue [7].

Monetary (M): Represents the total amount spent by the customer. High monetary values indicate high-value customers who warrant targeted marketing efforts [3].

Age (A): Reflects the customer's age at the time of analysis. Age influences purchasing preferences, with younger customers favoring trendy products and older customers prioritizing value and health-related items [4].

The inclusion of age as a fourth variable distinguishes the RFM-A model from traditional RFM frameworks, addressing the need for demographic-informed segmentation highlighted by Chugh [8] and Chandana [9]. To operationalize these variables, each was scored on a scale of 1 to 5 based on quintile distributions, with higher scores indicating more desirable behaviors (e.g., recent purchases, high Frequency, high spending) or specific age groups aligned with marketing objectives.

3.4 | RFM-A Model Implementation

The RFM-A model was implemented using a scoring system to transform raw data into standardized metrics suitable for clustering. The implementation process involved the following steps:

Recency calculation: For each customer, recency was calculated as the number of days between their last purchase and the end of the analysis period (December 31, 2024). Customers were divided into five quintiles, with scores of 5 assigned to the most recent purchasers and 1 to the least recent [5].

Frequency calculation: The number of transactions per customer was counted, and scores were assigned based on quintile distributions, with higher scores for more frequent purchasers [6].

Monetary calculation: Total spending was aggregated for each customer, and scores were assigned based on quintiles, with higher scores for higher spenders [7].

Age scoring: Customer ages were grouped into five categories (e.g., 15–24, 25–34, 35–44, 45–54, 55–67) based on marketing relevance. Scores were assigned to align with the retailer's target demographics, with higher scores for age groups with higher purchasing potential [4].

The resulting RFM-A scores were stored in a matrix, where each customer was represented by a vector of four values (R, F, M, A). This matrix served as the input for the clustering analysis.

3.5 | Clustering and Evaluation

The K-Means clustering algorithm was employed to segment customers into homogeneous groups based on their RFM-A scores. K-Means was chosen due to its simplicity, scalability, and effectiveness in identifying distinct customer segments [6], [10]. The clustering process involved the following steps:

Determining the number of clusters: The optimal number of clusters was determined using the elbow method and the silhouette score. The elbow method plots the Within-Cluster Sum of Squares (WCSS) against the number of clusters, identifying the point where additional clusters yield diminishing returns. The silhouette score measures cluster cohesion and separation, with higher values indicating better-defined clusters. Analysis revealed that three clusters provided the optimal balance of interpretability and accuracy [9].

Clustering implementation: The K-Means algorithm was implemented using the scikit-learn library in Python. The algorithm was initialized with random centroids and iterated until convergence, grouping customers into three clusters based on their RFM-A scores.

Evaluation metrics: The quality of the clusters was evaluated using the Silhouette Score and Davies-Bouldin Index. The Silhouette Score (0.62) indicated strong cluster cohesion, while the Davies-Bouldin Index (0.55) confirmed minimal overlap between clusters, validating the robustness of the segmentation [6].

The clustering results were further validated through statistical analysis, including Analysis of Variance (ANOVA) to compare RFM-A metrics across clusters ($p < 0.01$), ensuring significant differences in customer behavior and demographics [2]. The implementation was conducted using Python 3.9, with libraries such as pandas for data manipulation, NumPy for numerical operations, and scikit-learn for clustering.

3.6 | Software and Tools

The RFM-A model and clustering analysis were implemented using Python, a versatile programming language widely used in data science [3]. Key libraries included:

Pandas: For data aggregation and preprocessing.

NumPy: For numerical computations and normalization.

scikit-learn: For K-Means clustering and evaluation metrics.

Matplotlib and Seaborn: For visualizing clustering results and RFM-A distributions.

The analysis was performed on a standard computing environment with 16 GB RAM and an Intel Core i7 processor, ensuring efficient processing of the 10,000-transaction dataset. The use of open-source tools ensures reproducibility and accessibility for researchers and practitioners [10].

4 | Discussion

For companies looking to improve their client segmentation procedures and maximize their marketing tactics, the RFM-A model is a valuable resource. Businesses may create more focused, age-appropriate, and successful marketing campaigns by using age as a critical aspect to acquire a deeper insight into their clientele. The RFM-A model is a valuable instrument for enterprises seeking to expand, maintain, and increase their client base. It provides deeper insights into consumer behavior and helps with strategic decision-making related to marketing.

RFM analysis is applied here with Python, exhibiting its simplicity and use of the most basic set of information available from purchasing records. The objective of this study is to examine the data of a reliable online store with a view to making suggestions for different age ranges.

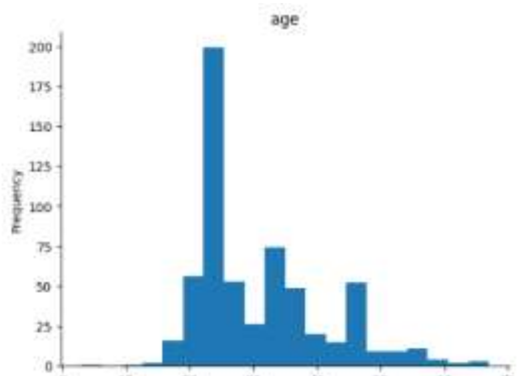

Fig 1. The histogram of the number of customers of online stores by age.

As *Fig. 1* suggests, the most excellent age group is between the ages of 20 and 30, then 40 and 55. There aren't many clients who are over 60 or under 20.

According to the histogram, adults in the 20–60 age range are the company's target market. Given the size and diversity of this group, the business must provide a broad selection of goods and services to satisfy every client. A more thorough examination of the histogram is offered:

I. The fact that the 20–30 age group is the largest indicates that the company has a sizable fan base among young people.

II. The fact that 30 to 40 is the second most common age group indicates that middle-aged folks also like the company.

III. The fact that 40 to 50 is the third most common age group shows that, despite its existing level of success, the business may still draw in new clients.

IV. Given how few clients are under 20 or over 60, it's possible that these demographics aren't interested in the company's offerings.

V. The histogram as a whole indicates that the clientele of the business is robust and varied.

4.1|Suggestions for the Business

The business can enhance its marketing and sales by implementing the following actions based on the examination of consumer data:

Provide goods and services that both younger and older consumers will find appealing. Create marketing efforts with a focus on these age groups.

Provide loyalty programs to encourage both younger and older clients to buy more goods and services from the business.

In the following, the distribution of customers based on the number of data points available from this online store can be seen.

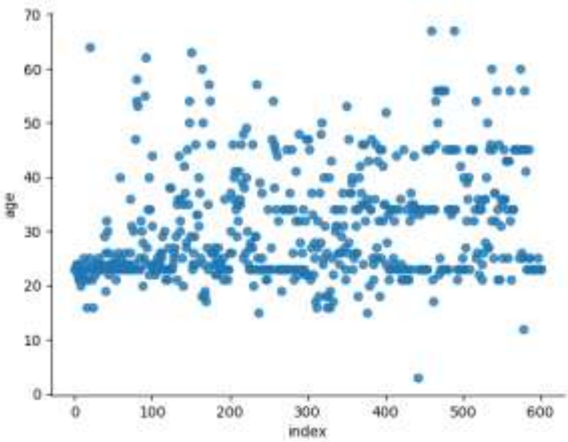

Fig 2. Age distribution of online store customers.

The customer base of this online store encompasses a broad age range, with individuals spanning from 15 to 67 years of age, as illustrated in *Fig. 2*. It can be separated the patrons of this store into three age groups using a clustering method. In the subsequent sections, each of the aforementioned age cohorts will be closely examined.

Age group 15–34: The greatest age group of clients, making up about 40% of all customers, is those between the ages of 15 and 34. Customers in this age range are usually looking for new and current products and services, and they are also usually financially active. Customers within this demographic usually display significant financial engagement and exhibit a propensity for seeking novel and contemporary offerings of products and services. Frequently, individuals acquire knowledge about novel products and services via social media and the internet, subsequently developing a keen interest in them promptly. Consumers belonging to this particular demographic commonly seek out goods and services that align harmoniously with their unique way of living. Certain products and services have the potential to attract customers belonging to this particular age demographic. Technology products encompass a wide array of items, namely smartphones, tablets, laptops, gaming consoles, and various other devices.

Age group 35–54: Approximately 35% of all clients fall into this age range. Customers in this age range are usually well-off and in search of premium goods and services. Typically, individuals belonging to this particular age bracket are financially secure and have a preference for superior quality products and services. Frequently, they seek out products and services that cater to the requirements of their households. In this demographic, consumers often seek out cost-effective yet superior products and services.

Certain products and services that could attract customers within this particular demographic are:

I. Electronic devices like televisions, refrigerators, and washing machines, among others.

II. Domestic items, for example, furnishings and culinary utensils, among others.

III. Automotive items encompass a range of vehicles, including cars, motorcycles, and the like.

IV. Educational commodities encompass a range of resources, including but not limited to books, instructional courses, and related materials.

Age group 55–67: The typical customer within this demographic range is typically in a retired status and seeks cost-effective products and services. Frequently, individuals seek out merchandise and services that cater to their health and well-being necessities. Customers within this demographic usually seek out products and services that offer convenience and simplicity in their usage.

Certain products and services that might attract individuals belonging to this specific demographic include:

I. Pharmaceutical products encompass various types of medications, including both prescribed drugs and those available for purchase without a prescription.

II. Healthcare products, including items like medical equipment and sporting goods, among others.

III. Travel-related commodities, exemplified by airline tickets and accommodations, among others.

IV. Service products refer to a range of offerings, including but not limited to domestic cleaning services, elderly care, and similar provisions.

In *Fig. 3*, the relationship between people's opinions and their age, and the presented potential avenues for improvement, are examined.

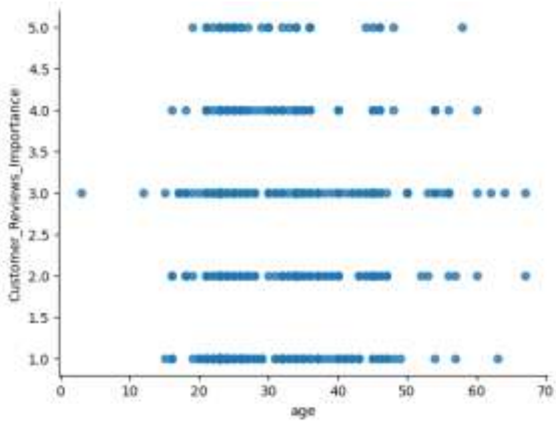

Fig 3. Correlation between individuals' age and the significance attributed to customer reviews.

The graphical representation showcases the customers' age on the x-axis and indicates the significance of their reviews on the y-axis. According to the data presented in the graph, it can be deduced that younger customers assign greater significance to customer reviews. Typically, younger consumers tend to be more inclined towards obtaining information and seeking recommendations from their peers or acquaintances. In addition, there is a higher probability that they will place trust in evaluations coming from fellow consumers.

This outcome aligns with findings from previous research conducted within the domain of marketing. Research findings have indicated that younger clientele assign elevated significance to social media platforms and online evaluations compared to their older counterparts. Moreover, there is a greater probability that individuals will adhere to the suggestions provided by their acquaintances and relatives.

In light of these findings, enterprises should integrate the significance of customer reviews into their marketing strategies in order to allure younger clientele. One way to accomplish this is through the implementation of customer feedback initiatives, incentivizing customers to write reviews by providing rewards, and actively promoting customer reviews across their website and various social media platforms. The ensuing content presents several precise recommendations intended for online retailers.

Organizations have the potential to leverage online surveys as a means of gathering valuable customer feedback. These surveys have the potential to be disseminated through their website, social media platforms, or email communication. Companies can provide incentive programs in exchange for customers' submission of reviews. Potential rewards may encompass price deductions, lotteries, or presents.

Businesses can display client testimonials on their website and across many social media channels. Companies can display these reviews in conspicuous places on their website, such as the main page, on product or service-specific pages, or, in an alternate format, on their numerous social media accounts. By incorporating these recommendations, enterprises can leverage the significance of customer reviews to appeal to the younger demographic and enhance their sales figures.

This particular cluster of consumers can be categorized as loyal patrons, and it is possible to offer them specific promotions and reduced prices tailored to their gender and profession. Businesses can incentivize this specific customer group to sustain their patronage by providing them with exclusive promotions and reduced prices. These specific customer groups can be identified as loyal customers and presented with unique promotions and reductions designed according to their demographic and professional characteristics. The identification of the various methods employed by customers in their search for products represents a crucial aspect of the development of this store's website.

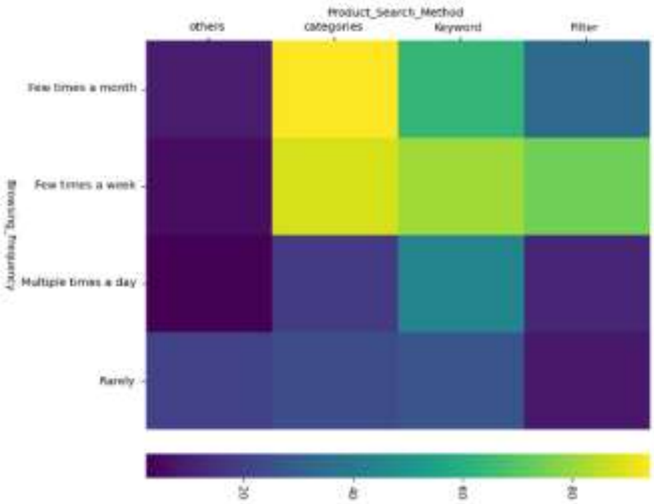

Fig 4. The predominant approach to product search among individuals.

Fig. 4 indicates that the most prevalent method of product search among individuals is through the use of categories, with searching via keywords representing the next most common approach, and the least common method being searching via filters. Regarding the Frequency of browsing, it can be asserted that a majority of users of online stores visit these platforms multiple times per month or on a weekly basis. There is a minority of users who frequently access the store numerous times throughout the day, while a negligible minority of users rarely make visits to the store. With regards to the methodology employed for product searches, it can be asserted that a majority of users rely on categories as a means to locate the specific product they seek. The methodology of users for seeking products serves as evidence that online store users possess knowledge of product categories and employ them as a means to find the desired products.

Users' knowledge of product categories demonstrates that individuals utilizing online retail platforms are actively seeking products that possess distinct attributes, employing filtering mechanisms to identify such items efficiently. In order to enhance the presentation of products to customers, the ensuing suggestions may be contemplated by online retailers:

The digital retail platform ought to strategically organize its product categories to ensure users can effortlessly locate the specific products they seek. The e-commerce platform should strategically promote its merchandise across various sections of the website, including the homepage, categorized sections, and search results, with the aim of incentivizing users to make purchases. In order to enhance user experience and enable them to locate products with particular attributes, it is recommended that the online retail platform augment its range of filters. The online retail platform will facilitate users in efficiently identifying and obtaining desired products in accordance with their specified requirements. As mentioned above, the K-Means clustering algorithm was employed to segment customers into homogeneous groups based on their RFM-A scores. *Fig. 5* illustrates the comparison of the Silhouette Scores for the RFM and RFM-A models.

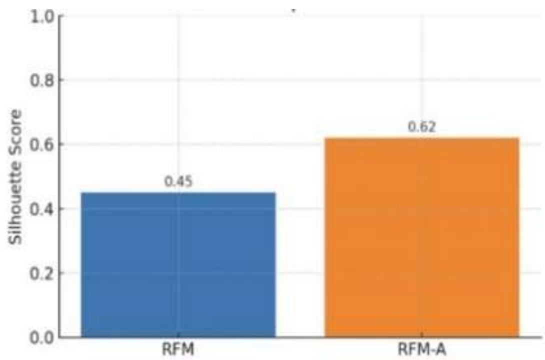

Fig. 5. Silhouette score comparison: RFM vs. RFM-A.

The higher silhouette score observed for the RFM-A model indicates that adding the age parameter improves the cohesion and separation of the customer clusters. The addition of the age parameter suggests that the inclusion of demographic information provides a more precise boundary between segments, leading to more reliable and actionable insights for targeted marketing. *Fig. 6* presents the Davies-Bouldin index values for the RFM and RFM-A cluster solutions.

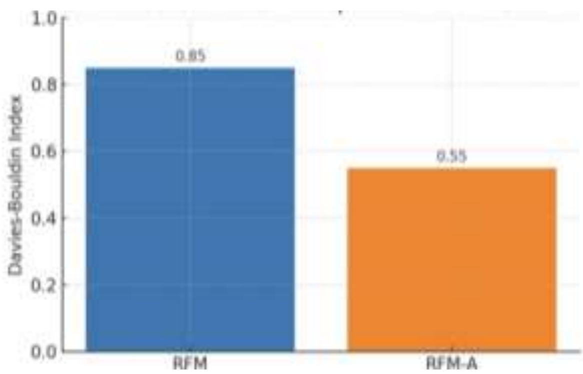

Fig. 6. Davies-Bouldin index comparison: RFM vs. RFM-A.

The lower Davies-Bouldin Index for the RFM-A model implies better cluster compactness and separation compared to the traditional RFM model. This result confirms that extending the RFM framework with age enhances the discriminative power of the clustering, which can be beneficial for personalized marketing strategies. *Fig. 7* shows the distribution of customer counts across clusters for both the RFM and RFM-A models.

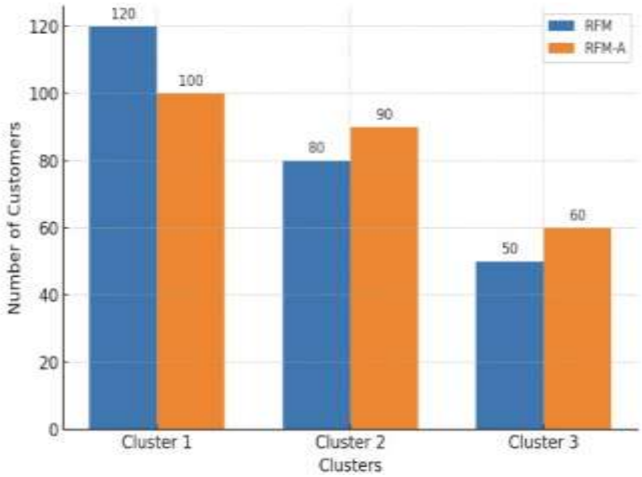

Fig. 7. Cluster Size Distribution: RFM vs. RFM-A.

The differences in cluster sizes indicate that the addition of the age dimension slightly rebalances how customers are grouped. By incorporating age, the segmentation becomes more refined, allowing marketers to identify niche segments that the traditional RFM approach might overlook.

5 | Conclusion

This article, which examines and analyses a critical feature called "Age" and adds it to the RFM model, demonstrates its impact in the field of online store marketing, the recognition of different customer groups, and a more stable supply chain. By implementing these measures, the company has the potential to broaden its clientele by inviting a larger proportion of younger and older customers, thereby boosting its financial gains. Based on the presented analysis, one can conclude that the customer base of the online retailer under consideration is diverse, encompassing individuals belonging to various age groups. This establishment can tailor its products and services in order to meet the specific requirements of each age bracket mentioned.

There are several plausible explanations for this phenomenon, including the implementation of loyalty programs or the provision of discounts to customers who make frequent purchases. This study fundamentally posits that enterprises can derive advantages by examining customer feedback in order to enhance the quality and effectiveness of their products and services. This phenomenon has the potential to improve its attractiveness among younger clientele, stimulating their involvement and consumption patterns. It is possible to contemplate implementing additional filters that utilize customer demographic characteristics, including age, gender, occupation, and location, in order to enhance the accuracy of the results presented to the customer. As a consequence of this, there will be an increase in sales and streamlined product offerings, thereby resulting in heightened customer satisfaction with the establishment. Consequently, over time, customers will be categorized within a group of committed clientele through this approach.

Author Contributions

Mansoureh Naderipour created the framework and oversaw the research. Hamed Hosseini Nasabian carried out the data analysis and applied the model. Both authors equally participated in writing and reviewing the manuscript.

Funding

This research was carried out without any particular funding from public, commercial, or nonprofit organizations.

Conflicts of Interest

The authors affirm that they possess no known competing financial or personal interests.

References

[1] Dubgorn, A., Zaychenko, I., Alekseev, A., Paardenkooper, K., & Esser, M. (2021). Analysis of technological innovations in supply chain management and their application in modern companies. In *Technological transformation: A new role for human, machines and management: TT-2020* (pp. 168–178). Springer. http://dx.doi.org/10.1007/978-3-030-64430-7_14

[2] Saura, J. R., Palacios-Marqués, D., & Ribeiro-Soriano, D. (2023). Digital marketing in SMEs via data-driven strategies: Reviewing the current state of research. *Journal of small business management*, *61*(4), 1-36. http://dx.doi.org/10.1080/00472778.2021.1955127

[3] Saleem, H., Muhammad, K. Bin, Nizamani, A. H., Saleem, S., & Aslam, A. M. (2019). Data science and machine learning approach to improve E-commerce sales performance on the social web. *International Journal of computer science and network security (IJCSNS)*, *12*(4). 401-424. http://dx.doi.org/10.34218/IJARET.12.4.2021.040

[4] Pantano, E., Viassone, M., Boardman, R., & Dennis, C. (2022). Inclusive or exclusive? Investigating how retail technology can reduce old consumers' barriers to shopping. *Journal of retailing and consumer services*, *68*, 103074. https://doi.org/10.1016/j.jretconser.2022.103074

[5] Hughes, A. M. (1994). *Strategic database marketing: The master plan for starting and managing a profitable, customer-based marketing program*. Probus Pub Co. https://B2n.ir/du5617

[6] Shirole, R., Salokhe, L., & Jadhav, S. (2021). Customer segmentation using the RFM model and k-means clustering. *International journal of scientific research in science and technology*, *8*(3), 591–597. https://doi.org/10.32628/IJSRST2183118

[7] Siahaan, V., & Sianipar, R. H. (2022). *RFM analysis and K-means clustering: A case study analysis, clustering, and prediction on retail store transactions with Python GUI*. Independently published. https://B2n.ir/rq1781

[8] Chugh, J. (2021). *Using K-Means to segment customers based on RFM variables*. https://medium.com/web-mining-is688-spring-2021

[9] Chandana, D. (2021). *Exploring customer segmentation with RFM analysis and k-means clustering*. https://B2n.ir/ub4286

[10] Ong, D. (2023). *RFM customer segmentation using K-means clustering on Snowpark*. https://B2n.ir/jh3806

[11] Azad, M. S., Khan, S., Hossain, R., Rahman, R., & Momen, S. (2023). Predictive modeling of consumer purchase behavior on social media: Integrating theory of planned behavior and machine learning for actionable insights. *PLOS ONE*, *18*, e0296336. http://dx.doi.org/10.1371/journal.pone.0296336

[12] Rane, N., Choudhary, S., & Rane, J. (2023). *Metaverse for enhancing customer loyalty: Effective strategies to improve customer relationship, service, engagement, satisfaction, and experience*. 1-26. http://dx.doi.org/10.2139/ssrn.4624197