Advances in Industrial Engineering, December 2025, 59(2): 343-358
DOI: 10.22059/aie.2025.394483.1943

RESEARCH PAPER

# Identifying Customer Segments in E-Commerce: A Data-Driven Framework Using Transactional Patterns

**Mehran Akbarpour[1], Ali Bozorgi-Amiri[2*] and Mostafa Setak[3]**
[1]*M.Sc. Student, Department of Industrial Engineering, K. N. Toosi University of Technology (KNTU), Tehran, Iran.*
[2]*Associate Professor, School of Industrial Engineering, College of Engineering, University of Tehran, Tehran, Iran.*
[3]*Associate Professor, Department of Industrial Engineering, K. N. Toosi University of Technology (KNTU), Tehran, Iran.*

Received: 1 May 2025, Revised: 1 August 2025, Accepted: 11 August 2025


## Abstract

This study proposes a structured and interpretable approach to customer segmentation by enhancing the classical RFM (Recency, Frequency, Monetary) model. The research addresses the common limitation of traditional RFM-based methods, which often overlook behavioral diversity and lack flexibility for practical, data-driven decision-making. The main objective is to develop a segmentation framework that provides actionable insights based on transparent, rule-based logic rather than opaque clustering algorithms. Using transactional data from a leading e-commerce platform in the Netherlands, the methodology applies quartile-based scoring to RFM indicators and maps customers into six distinct behavioral segments. Visual analytics are employed to support interpretation, enabling businesses to tailor engagement strategies for each group. The results demonstrate improved segment differentiation, managerial interpretability, and relevance to real-world applications. The study also highlights the model's adaptability to other customer-centric industries, with future research directions focused on incorporating machine learning and behavioral enrichment for Customer Lifetime Value (CLV) prediction.

**Keywords:** Customer Segmentation, RFM Model, Behavioral Analysis, Online Retail, Personalized Marketing, Data Driven Analysis.

## Introduction

In today's highly competitive and data-driven retail landscape, understanding customer behavior is critical for businesses aiming to enhance engagement, retention, and profitability. One of the most effective ways to achieve this is through customer segmentation, which allows firms to tailor strategies to the specific needs and behaviors of different customer groups. However, conventional segmentation methods, particularly traditional implementations of the RFM (Recency, Frequency, Monetary) model, often fall short in capturing the diversity and complexity of customer interactions. These limitations result in generalized insights that may not translate effectively into action. Therefore, there is a pressing need for more precise, interpretable, and scalable segmentation approaches that reflect real behavioral patterns and support practical decision-making. This study addresses that gap by proposing a structured RFM-based framework designed to generate clear, data-driven customer segments aligned with

* Corresponding author: (Ali Bozorgi-Amiri)
Email: alibozorgi@ut.ac.ir

business objectives.

For instance, the choice or quantity of items purchased may be affected by whether one is shopping alone, with someone, or with parents or a spouse. The use of technology in-store behavior, including behavioral methods and their applications in retail, is detailed in [1]. Beyond conventional methods, research has been undertaken to push the boundaries of segmentation strategies, adopting approaches tailored to the nuances of the available data. The intricacies of this unique segmentation are navigated, shedding light on the insights gleaned and the potential it holds for a more targeted and effective approach to customer engagement.

While the RFM model has been a cornerstone in numerous research endeavors, this article takes a distinct approach, marked by a commitment to unrivaled granularity. In the expansive realm of retail research, the model has traditionally been applied broadly, offering valuable insights at an aggregate level. However, in this article, sophistication is elevated by venturing into the uncharted territory of individualized segmentation [2]. The approach represents a paradigm shift in the application of the RFM model, redefining its usage to intricately dissect individual customer research within the context of a prominent retail store in the Netherlands [3].

What distinguishes this study from previous RFM-based segmentation research is its focus on combining statistical transparency with practical interpretability. Rather than relying on clustering algorithms that may require parameter tuning and post hoc interpretation, this study introduces a rule-based, quartile-driven segmentation method that maps RFM scores directly to distinct behavioral categories. This approach ensures that each segment is both data-grounded and intuitively meaningful, allowing businesses to trace decisions back to specific customer traits. By emphasizing clarity, reproducibility, and business alignment, the proposed framework offers a novel contribution to customer analytics—bridging the gap between technical rigor and practical application.

This study introduces an innovative approach. The objective is clear: to move beyond conventional segmentation approaches and tailor insights to the idiosyncrasies of each customer. By personalizing the RFM model, the limitations of generic categorizations are transcended, delving deep into the unique intricacies shaping the purchasing behavior of individual consumers. The theoretical framework is not the final stop; The study proceeds with an in-depth examination of the preprocessing and modeling steps of data preprocessing. Recognizing the importance of pristine data, the dataset is meticulously cleaned and transformed to ensure the accuracy and reliability of subsequent analyses. The research begins with refining raw data, addressing outliers, and normalizing variables to create a solid foundation.

Building on this foundation, a suite of data mining techniques is employed to extract meaningful patterns and insights. Each step is a carefully crafted progression towards uncovering hidden gems within the dataset. From clustering algorithms to association rule mining, the arsenal of tools is tailored to the specific characteristics of the retail landscape, offering a comprehensive understanding of customer behavior. As these step-by-step processes are navigated, the objective is not only to implement the RFM model but to enhance its capabilities through the integration of cutting-edge data mining techniques. The aim is to transcend traditional segmentation approaches and capture the intricacies of customer interactions, enabling the retail giant to make informed decisions resonating with its diverse customer base.

This study systematically explores the stages of data preprocessing and data mining, where each analytical stride brings us closer to systematically analyzing customer segmentation patterns in the dynamic retail environment of the Netherlands [4]. This section explores the complexities that are demystified, insights gained are shared, and the way is paved for a more targeted and personalized approach to customer-centric strategies. Following the foundational analysis with an exhaustive exploration of data preprocessing and cutting-edge data mining techniques, the focus now turns to the heart of this article – the revelation of distinct customer segments based on the robust RFM model. Through meticulous implementation, a spectrum of customer archetypes is unearthed within the dataset of the esteemed retail establishment in the Netherlands [5].

Lastly, our study presents an inventive segmentation approach for understanding the position of customers based on their recent behavior. This approach provides a more granular alternative to traditional segmentations, allowing for a more nuanced and granular understanding of customer segments. By adopting this innovative segmentation method, we aim to unravel the intricacies of customer behavior, providing businesses with targeted insights for personalized strategies.

## Literature Review

Given the research focus on the RFM model and recent developments in customer segmentation, a preliminary literature review was conducted to establish a foundation for the research problem. Recent articles within this field have contributed to shaping contemporary perspectives on customer segmentation. The intent is to delve deeper into the existing literature, particularly focusing on the RFM model and customer segmentation, to gain a more nuanced understanding of these critical aspects. This targeted literature review will form the cornerstone of the research, providing essential insights and contextualizing the study within the broader landscape of relevant scholarship

### Overview of the RFM Model

An inherent strength of the e-commerce landscape lies in its ability to trace and retrieve customer shopping behaviors, facilitated by either internal web analytics tools or external web crawling software [6]. The abundance of online shopping information, encompassing browsing habits and purchasing patterns, has positioned behavioral segmentation as the predominant approach to customer categorization [7].

Typically, RFM data are initially derived from historical purchasing data and subsequently mapped into discrete groups, a process known as RFM scoring. Two prevalent methods for categorizing customers into scored groups are the customer quintile and behavior quintile scoring methods. Customer quintile scoring involves sorting customers in descending Frequency and Monetary orders or in ascending Recency order, breaking them into five equal groups or quintiles. For example, and assigned the top quintile a score of five, the next quintile a score of four, and so forth. In contrast, behavior quintile scoring organizes customers into five quintiles based on a behavioral measure, potentially resulting in varying numbers of customers in each quintile. [8]. For instance, assigned a score of "1" to customers who shopped only once and a score of "5" to those who shopped over six times. Assigned a score of "1" to credit-card users with transactions less than $50 million and a score of "5" to those with transactions exceeding $300 million. Typically, customers who make the most recent, frequent, and highest-value purchases are assigned the highest scores in each of the RFM dimensions, respectively. Following data mapping, clustering algorithms can automatically categorize the RFM data into specific clusters [9].

The researchers not only segmented customers but also calculated the CLV for each obtained segment. Emphasized the significance of customer segmentation as a crucial step in fostering and sustaining customer relationships, ultimately leading to increased sales [10]. The article proposed a market segmentation method for retailers based on customer lifespan, indicated by exchange volume, and a segmentation approach based on purchase history [11].

**Customer Segmentation**

The customer segmentation process has three main components: segmentation variables, method, and validation. The relationship between product decision-making information systems and big data-driven is inspired by [12]

Customer segmentation is a foundational strategy in marketing that enables businesses to tailor their approaches to diverse customer needs, tracing its roots to Smith's definition in 1956, aiming to accommodate the heterogeneity in demand and tailor marketing strategies accordingly. The retail literature has exhibited a pronounced interest in segmentation studies, with early works focusing on offline customer segmentation and subsequent studies delving into online customer segmentation. As the retail landscape evolves, attention has shifted towards delineating segments of customers navigating various channels. Studies on multichannel segmentation typically rely on either self-reported (survey) data or behavioral data, with a few exceptions integrating both approaches [13]. Table 2 succinctly outlines the characteristics of key multichannel customer segmentation studies employing clustering algorithms to unveil behavioral heterogeneity.

While studies based on declared channel preferences may suffer from common method bias, employing behavioral data, as highlighted by [14]. Provides a more actionable approach. For instance, RFM variables are readily accessible in firms' databases [15].

In the realm of advanced analytics shaping customer-centric decision-making, RFM segmentation serves as a pivotal tool for customer value analysis and targeted outreach. This model offers a simplified scoring framework that captures customer behavior through transaction history, using pre-defined RFM indicators previously outlined. However, inherent limitations hinder RFM's maximal contribution to decision-making in customer-centric strategies. These include the neglect of key behavioral attributes, the use of inefficient scoring methods, and the lack of scientific integration among RFM dimensions [16].

Moreover, through an examination of recent studies focusing on customer segmentation utilizing the RFM model, this research endeavors to address these challenges by striving for a distinctive and precise segmentation approach [17].

In essence, the text describes a research initiative led by Ho and colleagues that thoroughly examines the RFM model, going beyond its traditional boundaries. The focus extends to understanding how the model has evolved and expanded, and how each of its components contributes to the overall understanding of customer behavior. The schematic diagram visually represents this comprehensive exploration.

In Figure 1, a comprehensive discussion and visual representation of the research conducted on the RFM model are presented. This figure provides an in-depth overview, offering detailed insights into the methodologies, findings, and visualizations associated with the application of the RFM model in customer segmentation. The visualizations in Figure 1 aim to enhance the clarity and accessibility of the RFM model's exploration, making it an indispensable reference for researchers and practitioners interested in leveraging this model for effective customer segmentation.

In Table 1, we present a comprehensive examination of the significance of customer segmentation, exploring the various channels through which this segmentation has been implemented [18]. This analysis encompasses a summary of findings from multiple studies, offering a consolidated overview of the importance of customer segmentation and the diverse strategies and methods employed in its application. The table serves as a valuable resource for stakeholders, researchers, and decision-makers seeking quick and concise insights into the nuanced landscape of customer segmentation in different contexts [19].

The problem at hand revolves around the need for effective customer segmentation in the dynamic retail environment. As outlined in the preceding content, the objective is to identify distinct categories of customers based on their behavior, utilizing the RFM model and innovative data mining approaches. The overarching goal is to gain a comprehensive understanding of customer segments, enabling personalized and targeted strategies for engagement [20].

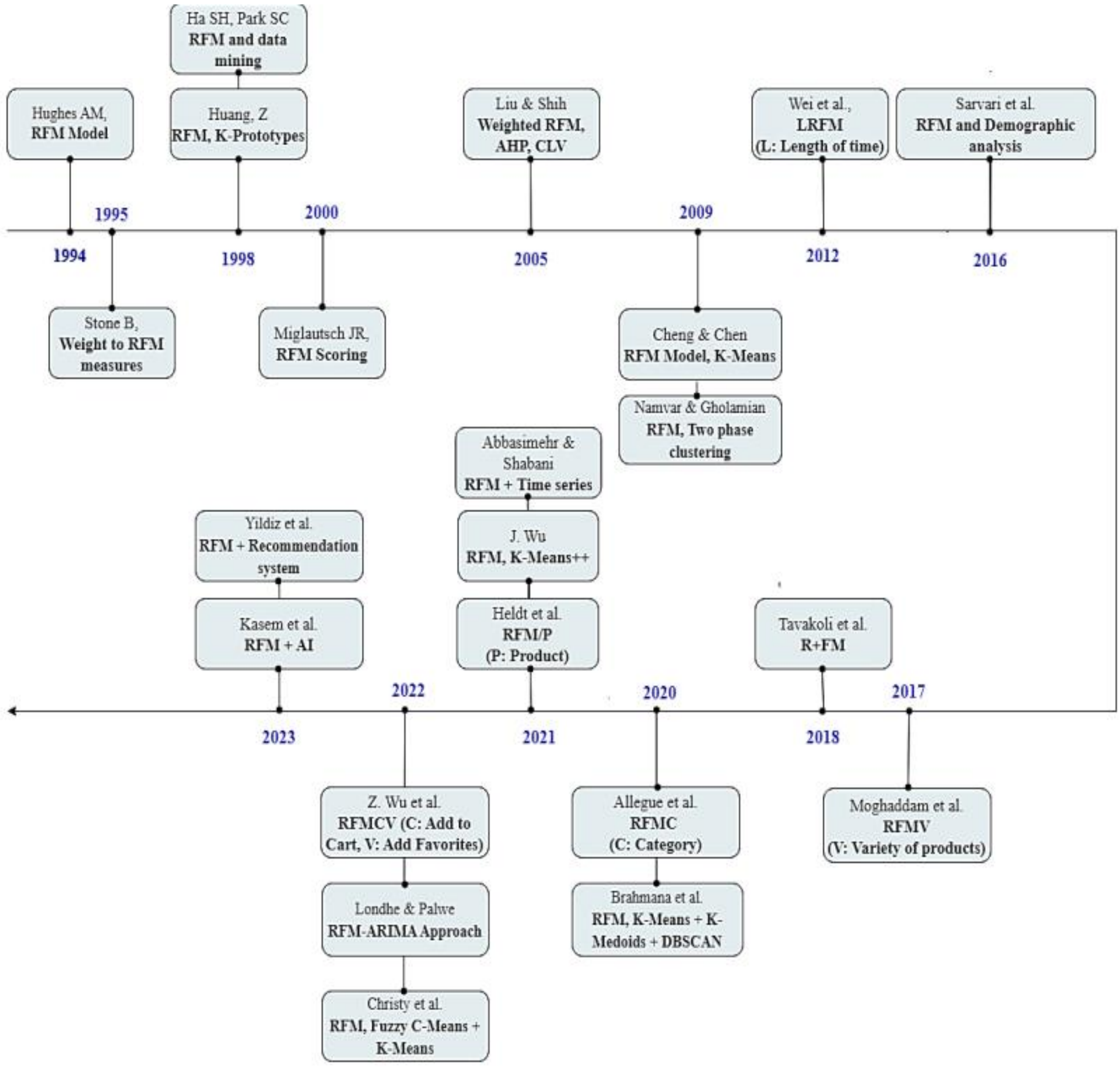


**Figure 1**. Overview of enhancements in the RFM model (Ho et al., 2023)

One key challenge is the high volume of available data, necessitating meticulous preprocessing to ensure its accuracy and relevance. Additionally, the content highlights the importance of focusing on a specific market, such as the United Kingdom, to tailor analyses to regional nuances. The ultimate aim is to go beyond traditional segmentation approaches and capture the intricacies of customer interactions. [21].

In comparison to recent research endeavors, our study makes substantial contributions in three key areas. Firstly, we place significant emphasis on the preprocessing of the dataset, ensuring that our data is meticulously cleaned and transformed for accuracy and reliability. This initial step lays a robust foundation for subsequent analyses, enhancing the overall quality of our study.

Secondly, our research introduces an innovative RFM model implementation. We go beyond the conventional applications of the RFM model and tailor its usage to the unique characteristics of our dataset and research objectives. This innovative approach enables us to glean deeper insights into customer behavior, transcending traditional boundaries in the realm of RFM model implementation.

**Table 1**. RFM-Based Analysis of Multichannel Impact: A Comprehensive Overview of Studies on Customer Segmentation

| Authors (Year) | Channel(s) | Product category (ies) | Aim of study | Data prior to online channel introduction | RFM segmenta-tion |
|---|---|---|---|---|---|
| **Gül & Şen (2022)** | Mobile App Push Notifications | Retail (Turkey) | Design effective push notifications based on RFM segmentation | No | ✓ |
| **Wong et al. (2024)** | Online Retail | E-commerce | Compare RFM with clustering techniques for customer segmentation. | No | ✓ |
| **Wang (2025)** | Digital Marketing | Durables and apparel | Investigate channel migration and channel Loyalty | No | ✓ |
| **Kushwaha and Shankar (2013)** | Catalog and online Channel | 22 non-grocery categories | Compare the monetary value of single- and multichannel customers | No | No |
| **Hernant and Rosengren (2017)** | Online channel added to store channel | Non-grocery utilitarian products | Assess the sales effect of adding an online Channel | ✓ | No |
| **Pauwels and Neslin (2015)** | Store channel added to catalog and online channel | Durables and apparel | Assess the sales effect of adding an offline channel | ✓ | No |
| **Melis et al. (2016)** | Online channel added to store channel | Groceries | Investigate whether online channel adoption increases share of wallet | No | No |
| **Montaguti et al. (2016)** | Stores, mail-order, phone and online channel | Books | Investigate if a marketing campaign increases multichannel shopping and profitability | No | No |
| **Kumar et al. (2018)** | Stores and online | Alcoholic beverages | Analyse effects of multichannel shopping on spending, frequency, and profitability | No | No |

## Method

As we delve into the intricacies of our endeavors, Figure 2 stands as a crucial companion, unraveling the specifics of the actions we've undertaken. In this detailed exploration, we embark on step-by-step research, deciphering the nuances and intricacies of the processes illustrated in Figure 1. This visual guide is designed to provide clarity and depth, offering a comprehensive understanding of each sequential action. The methodological process is systematically detailed as elucidated in Figure 2. This figure provides a structured depiction of the methodological approach, outlining the specific steps involved in the process. Refer to Figure 2 for a comprehensive understanding of the sequence and details behind the actions discussed.

The data collection process for this study involved sourcing information from an extensive online customer dataset within the Netherlands. The selection of this dataset was deliberate, driven by its relevance and strong correlation with the distinctive features of the RFM model. The online customer data set provides a rich and dynamic source, capturing a diverse array of customer interactions and transactions within the context of a prominent retail environment.

The choice of the Netherlands as the study's geographical focus ensures a localized and contextualized perspective, allowing for insights that are reflective of the unique characteristics of the Dutch market. Leveraging an online customer data network facilitates the exploration of real-time and historical interactions, offering a nuanced understanding of customer behavior in the digital retail landscape.

In alignment with the RFM model, this study applies a quartile-based implementation of the RFM model to construct interpretable customer segments based on behavior. This methodological approach allows for the extraction of meaningful patterns and insights that are instrumental in the subsequent application of data mining techniques. The integration of this dataset with the RFM model not only aligns with established principles of customer segmentation but also sets the stage for a comprehensive analysis, contributing to a deeper understanding of customer behavior in the Netherlands' retail landscape [22].

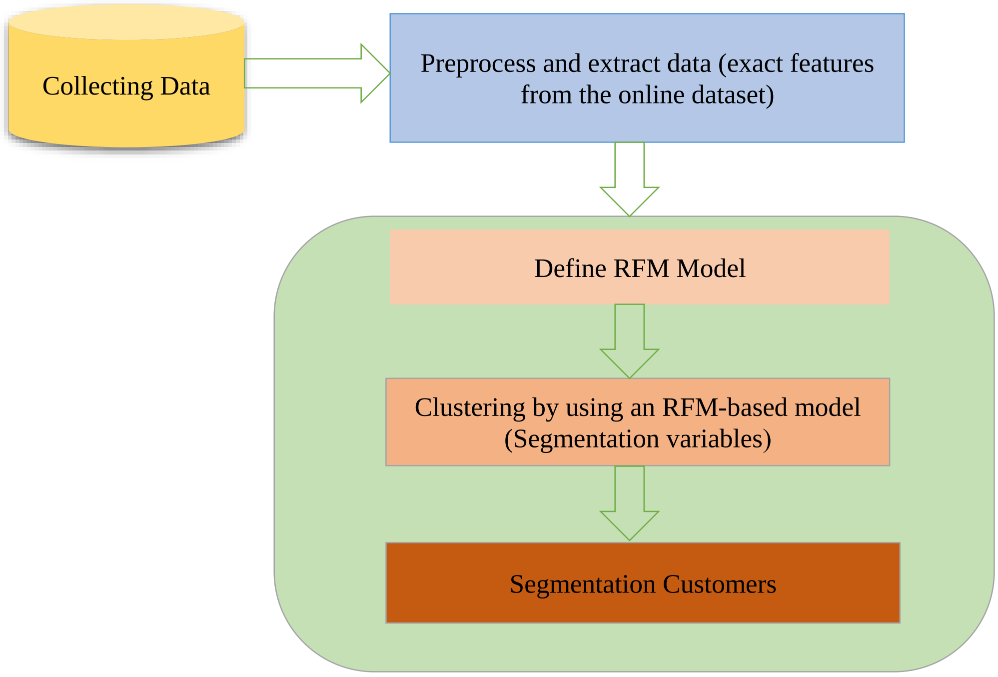


**Figure 2**. From Data Preprocessing to Individualized RFM Segmentation

**Table 2**. Online Retail Transaction Details

| InvoiceNo | StockCode | Description | Quantity | InvoiceDate | UnitPrice | CustomerID | Country |
|---|---|---|---|---|---|---|---|
| **536365** | 85123A | WHITE HANGING HEART T-LIGHT HOLDER | 6 | 12/1/2010 8:26 | 2.55 | 17850 | United Kingdom |
| **536365** | 71053 | WHITE METAL LANTERN | 6 | 12/1/2010 8:26 | 3.39 | 17850 | United Kingdom |
| **536365** | 84406B | CREAM CUPID HEARTS COAT HANGER | 8 | 12/1/2010 8:26 | 2.75 | 17850 | United Kingdom |
| **536365** | 84029G | KNITTED UNION FLAG HOT WATER BOTTLE | 6 | 12/1/2010 8:26 | 3.39 | 17850 | United Kingdom |
| **536365** | 84029E | RED WOOLLY HOTTIE WHITE HEART. | 6 | 12/1/2010 8:26 | 3.39 | 17850 | United Kingdom |

In line with our methodological approach, which emphasizes meticulous data preprocessing to ensure the accuracy and reliability of our analyses, the initial step involved managing the vast volume of available data. As delineated in Table 2, it became apparent that the United

Kingdom constituted a substantial portion of the dataset compared to other regions. Recognizing the importance of focusing on a specific market for in-depth analysis, the first crucial step in data preparation was to narrow down the transactional data exclusively to the United Kingdom. This strategic decision not only streamlined the dataset but also allowed us to delve deeper into the nuances of customer behavior within this specific geographical context.

By isolating the transactions within the United Kingdom, we aimed to gain a more granular understanding of customer interactions and purchasing patterns within this significant market segment. This focused approach is essential for extracting meaningful insights and ensuring that our subsequent analyses accurately capture the intricacies of customer behavior specific to the United Kingdom. The concentration on a specific geographical region enhances the relevance and applicability of our findings, providing a solid foundation for the subsequent stages of our analysis.

The data preprocessing research continues beyond this initial narrowing down of the dataset. Subsequent steps involve addressing outliers, normalizing variables, and refining raw data to create a robust foundation for our exploration. Each preprocessing step is carefully executed to pave the way for effective data mining techniques and the implementation of the RFM model, ultimately enabling us to unravel valuable insights into customer segmentation and behavior within the dynamic retail landscape.

In the methodological framework of our study, data preprocessing emerged as a pivotal step to ensure the quality and reliability of the dataset. The initial phase involved meticulous cleaning, where we addressed various aspects such as missing values, outliers, and inconsistencies within the data. This process aimed to enhance the overall integrity of the dataset, laying the groundwork for subsequent analyses. Through systematic data cleaning, we aimed to eliminate noise and discrepancies that could potentially impact the accuracy of our results.

Upon completing the data cleaning phase, we achieved a substantial volume of records, reflecting a refined and well-structured dataset. This post-cleaning dataset is characterized by its suitability for the implementation of the model. The decision to embark on data preprocessing was driven by the recognition of its critical role in facilitating meaningful insights from the subsequent application of the model. The refined dataset provides a solid foundation for uncovering patterns and trends within customer behavior, allowing for a more nuanced exploration of distinct customer segments. The availability of a significant and cleanse dataset positions us favorably for the model's implementation. The richness of the data ensures that the model analysis can be conducted with a high degree of accuracy, enabling us to derive valuable insights into customer segments. This methodological approach aligns with our overarching goal of pushing the boundaries of conventional segmentation strategies and adopting innovative techniques to unravel the complexities of customer interactions within the dynamic retail landscape of the Netherlands [20].

Building on the foundations laid by the data preprocessing phase, the next step in our analytical research involved the implementation of the model. The primary objectives of the RFM model were strategically aligned to extract nuanced insights from the dataset. Recency focused on capturing the latest micro-dates of customer transactions, shedding light on the temporal aspect of purchasing behavior. Frequency aimed to discern patterns in the repetition of customer purchases, providing valuable information on customer loyalty and engagement [24]. Lastly, Monetary delved into quantifying the purchase volume in terms of monetary value, offering a perspective on the economic significance of each transaction [25]. This multi-faceted approach within the RFM model allowed for a comprehensive analysis of customer behavior, enriching our understanding of the diverse facets influencing purchasing decisions [23].

As we delved into the implementation of the model, the significance of each model parameter became evident, showcasing its role in unraveling distinctive characteristics of customer segments. The results were subsequently organized and presented in Table 3, providing a structured and detailed overview of the segmented customer data. This tabular representation serves as a visual aid to better comprehend the intricate patterns and trends identified through the model, offering a comprehensive snapshot of the diverse customer segments within the retail landscape. The integration of modeling marks a pivotal step in our analytical research, paving the way for a deeper exploration of customer behavior and segmentation strategies tailored to the unique dynamics of the dataset.

**Table 3**. Customer Segmentation Insights: Recency, Frequency, and Monetary Analysis by CustomerID

| CustomerID | Recency | Frequency | Monetary |
|---|---|---|---|
| 12747 | 109 | 5 | 191.85 |
| 12748 | 70 | 96 | 1054.43 |
| 12749 | 130 | 3 | 67.00 |
| 12820 | 74 | 1 | 15.00 |
| 12821 | 214 | 1 | 19.92 |

In the pursuit of a meticulous customer segmentation strategy, our next step involves a granular examination of each RFM index, unraveling insights based on distinct customer behaviors. To achieve precision in this division, we opted for a quartile-based approach, where each RFM indicator is systematically categorized into quartiles. By doing so, any given customer is assigned to a specific part based on the quartiles associated with each of the RFM indicators. This meticulous division ensures a fine-grained understanding of customer behavior, capturing nuances in the model's aspects. The division process culminates in the derivation of an RFM score for each customer, offering a comprehensive depiction of their position relative to the model indicators. This score provides a succinct summary of each customer's engagement with the retail establishment, reflecting their recency of purchase, frequency of transactions, and monetary contributions. The outcome of this segmentation strategy is presented in Table 4, showcasing the RFM scores for individual customers. This structured presentation facilitates a clear view of how each customer aligns with the RFM indicators, laying the foundation for targeted and personalized marketing approaches.

**Table 4**. RFM Score-based Customer Segmentation Overview

| CustomerID | Recency | Frequency | Monetary | R_Quartile | F_Quartile | M_Quartile | RFMScore |
|---|---|---|---|---|---|---|---|
| 12747 | 109 | 5 | 191.85 | 3 | 4 | 4 | 344 |
| 12748 | 70 | 96 | 1054.43 | 4 | 4 | 4 | 444 |
| 12749 | 130 | 3 | 67.00 | 2 | 3 | 3 | 233 |
| 12820 | 74 | 1 | 15.00 | 4 | 1 | 1 | 411 |
| 12821 | 214 | 1 | 19.92 | 1 | 1 | 2 | 112 |

Table 4 serves as a visual representation of the effectiveness of our RFM-based customer segmentation, offering a snapshot of the diverse customer landscape. Each entry in the table corresponds to a unique customer, providing insights into their recency, frequency, and monetary behaviors. This scoring system not only aids in understanding the distribution of customers across different segments but also forms the basis for crafting tailored marketing strategies that resonate with the specific needs and preferences of each segment. As we delve into the intricacies of this model score-based segmentation, the significance lies not just in the numerical values but in the actionable insights they unlock. This approach goes beyond conventional segmentation methods, offering a dynamic and nuanced understanding of customer engagement. The subsequent sections of our analysis will further explore the implications of these model scores, shedding light on the potential strategies for targeted customer engagement and retention in the competitive retail landscape.

The table provides a comprehensive overview of customer behavior. Each row represents a different customer, identified by their unique CustomerID.

Recency: This metric indicates how recently a customer made a purchase. It is measured in days, with lower values indicating more recent activity. For instance, customer 12747 made a purchase 109 days ago, while customer 12821 made a purchase 214 days ago.

Frequency: Frequency refers to the number of purchases made by each customer within a specified period. For example, customer 12748 made 96 purchases, indicating high engagement, while customer 12821 made only one purchase.

Monetary: This metric represents the total monetary value of purchases made by each customer. It provides insight into the spending habits of customers. Customer 12748 spent a substantial amount, totaling 1054.43 units, whereas customer 12749 spent a smaller amount, only 67.00 units.

After listing the raw vlues for each customer, quartiles are calculated for each metric. Quartiles divide the data into four equal parts, allowing for a better understanding of distribution and comparison between customers.

R_Quartile: This quartile represents the recency score. It categorizes customers based on how recently they made a purchase relative to others. A higher R_Quartile indicates a more recent purchase. For instance, customer 12749 has an R_Quartile of 2, meaning their purchase was more recent compared to others.

F_Quartile: The frequency quartile categorizes customers based on the frequency of their purchases. Higher F_Quartile values indicate more frequent purchases. For example, customer 12748 has an F_Quartile of 4, indicating they are among the most frequent purchasers.

M_Quartile: This quartile reflects the monetary value of purchases. A higher M_Quartile signifies a higher monetary value. For instance, customer 12748 has an M_Quartile of 4, indicating they have spent a significant amount compared to others.

Finally, the RFM Score is calculated based on the quartiles of each metric for each customer. It condenses the information from Recency, Frequency, and Monetary value into a single score, providing a holistic view of customer behavior. For example, customer 12747 has an RFM Score of 344, indicating moderate recency, frequency, and monetary value.

**Result**

Based on the quartile-based RFM scores, customers were categorized into six behaviorally distinct segments. The classification was derived by analyzing combinations of RFM model values, and interpreting their relative rankings across all customers. The logic for assigning each segment is as follows:

Loyal Customers: Individuals with high frequency and monetary scores, along with recent activity—representing a core group of consistently engaged and high-value buyers.

High-Value Customers: Customers who have contributed significantly in terms of monetary value, often combined with strong recency or frequency metrics.

Strategic Spenders: Buyers with regular purchasing frequency and moderate-to-high monetary value, suggesting intentional and pattern-based buying behavior.

Seasonal Shoppers: Customers who show periodic engagement with moderate frequency, but have less recent activity—often linked to seasonal campaigns or events.

Spent More Than Others: Those with infrequent purchases but relatively high spending in one or more transactions, indicating low frequency but impactful monetary contribution.

At-Risk Customers: Individuals with declining or low recency and frequency scores, suggesting reduced engagement or potential churn risk.

This classification strategy was guided by both the numerical distribution of model scores and an interpretive understanding of typical e-commerce customer behaviors. By applying this structured logic, the segmentation offers meaningful insights for retention strategies and personalized marketing interventions.

It's crucial to emphasize that this segmentation is rooted in the available data and represents an innovative perspective tailored to the specific business context. The subsequent Figure 3 provides a visual representation of the classification quantities and the distribution of customers within each category, employing both bar and bubble forms for a comprehensive display. This visual aid enhances the interpretability of the segmentation, allowing for a more nuanced understanding of each customer category and its distinctive attributes.

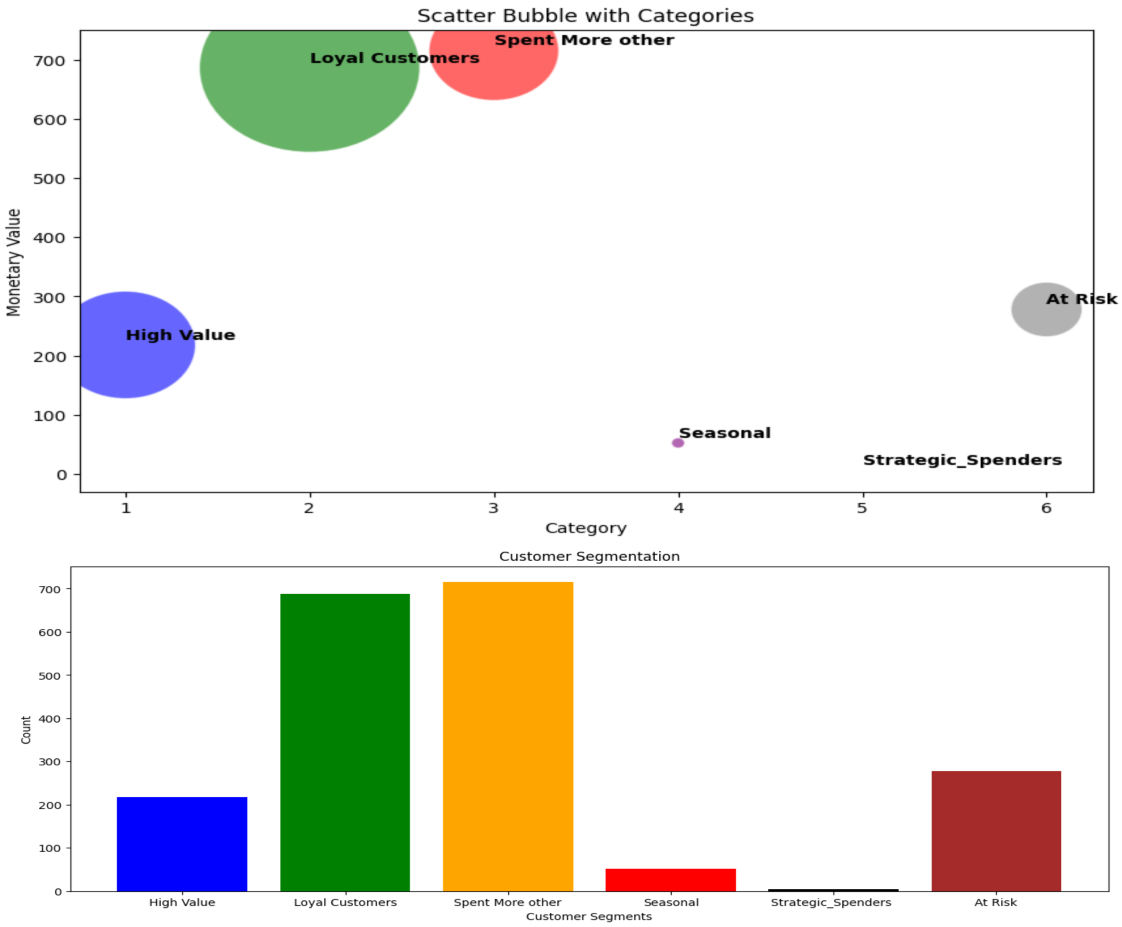


**Figure 3**. Visual Representation of Customer Segmentation: Bar and Bubble Charts Depicting Unique Categories

In Figure 4, a detailed breakdown of each customer category is presented, offering a comprehensive view of their respective indices. The diagram visually illustrates the distribution of RFM scores within each category, allowing for a nuanced understanding of how customers are distributed across the model indicators. This visualization serves as a valuable tool to assess the composition of each category, providing insights into the characteristics and behaviors of customers within each segment.

By showcasing the distribution of RFM indices within each category, Figure 4 aids in uncovering patterns and trends associated with customer behavior. This detailed breakdown facilitates a more nuanced analysis of customer segments, enabling businesses to tailor their strategies based on the distinct RFM characteristics exhibited by different customer categories. The visual representation not only enhances the clarity of the segmentation but also serves as a strategic guide for businesses seeking to optimize their engagement strategies and marketing efforts across various customer segments.

To enhance the clarity of our results, we employed visualization techniques. As part of this effort, Figure 5 is presented to depict the mid-range RFM values within each customer category, thoughtfully labeled for improved expression. This visualization aims to provide a more nuanced understanding of how customers are distributed across the mid-range of model's indicators within each identified category.

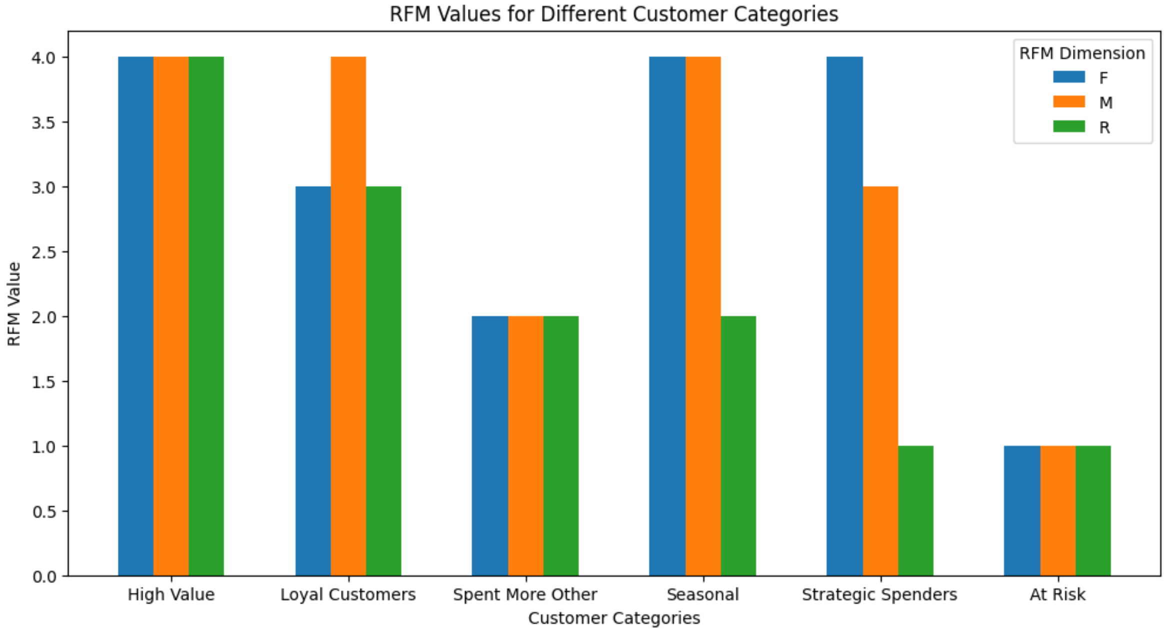


**Figure 4.** RFM Index Distribution Across Customer Categories

In Figure 6, we delve into exploring the relationships between individual indicators of the RFM model. This analysis aims to ascertain whether there is a discernible correlation between each RFM parameter separately. By examining this figure, we aim to understand the interplay between the model's aspects within the segmentation. This exploration is vital in determining if customers residing in different branches exhibit notable relationships in terms of their purchasing behavior.

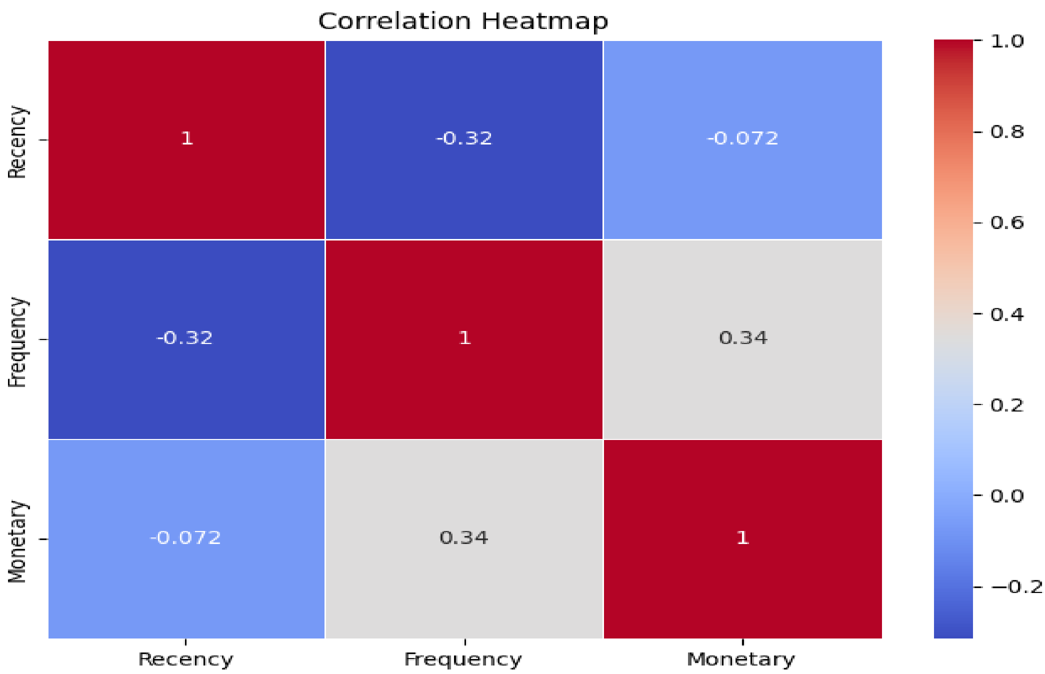


**Figure 5**. Correlation Heatmap

Each branch in Figure 6 represents a specific RFM indicator, and the intersections aim to uncover potential patterns or connections. This visual representation allows us to assess the degree of influence each model parameter holds within the identified customer segments. Analyzing these relationships provides valuable insights into the dynamics of customer behavior, helping us discern patterns that may guide targeted marketing efforts or tailored strategies for each segment.

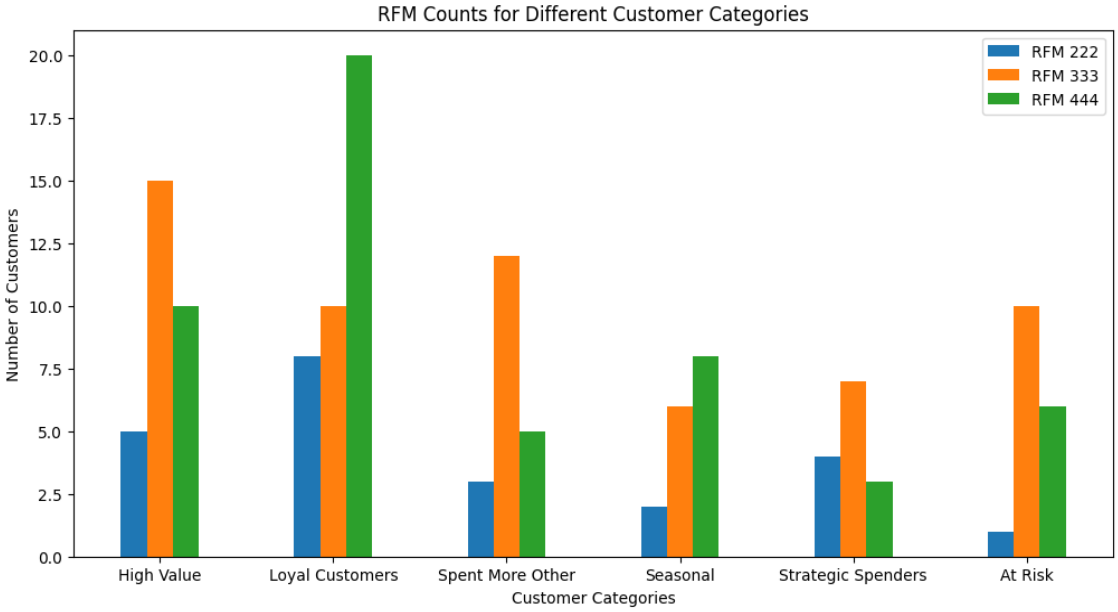


**Figure 6**. RFM Index Distribution in Customer Categories: Visualizing Mid-Range Insights

In Figure 6, each category is meticulously examined to reveal the specific RFM values that constitute the mid-range for customers falling into that category. This detailed breakdown allows for a more granular analysis of customer behavior within distinct segments. By providing a visual representation with clear labels, we aim to facilitate an intuitive comprehension of the mid-range RFM dynamics across different customer categories. This approach aids in identifying patterns, trends, and potential areas for strategic focus within each segment, contributing to a more insightful interpretation of the customer segmentation based on the model's indicators.

In this study, the process commenced with the collection of a robust dataset from an online customer network in the Netherlands. Due to the considerable volume of data and with a focus on the United Kingdom, which dominated the dataset, the initial step involved narrowing down the transactions exclusively to the United Kingdom. This targeted approach was undertaken to ensure a more accurate and relevant analysis of customer behavior within a specific market.

Following data collection, a meticulous pre-processing phase was executed to clean and refine the dataset. This step aimed to eliminate any inconsistencies or outliers, ensuring that the data was well-prepared for subsequent analysis. The United Kingdom-focused dataset was then subjected to pre-processing techniques to handle missing values, standardize formats, and address any other data quality issues. After this thorough pre-processing, the dataset was deemed suitable for implementing the RFM model.

The innovative implementation of the RFM model involved a detailed examination of Recency, Frequency, and Monetary indicators. The RFM scoring was performed by dividing customers into specific quartiles based on each indicator. This resulted in the assignment of RFM scores to each customer, providing a comprehensive understanding of their purchasing behavior. The final step involved visualization techniques, such as charts and diagrams, to present the categorized customers and their RFM scores. This visualization aimed to enhance comprehension and facilitate the identification of distinct customer segments, empowering businesses to tailor strategies based on individual customer behaviors

Understanding the category to which a customer belongs, especially if they continue to make transactions, is crucial for devising unique recovery strategies. By identifying the specific category a customer falls into based on their model indicators, businesses can tailor their recovery approaches to suit individual customer behaviors. For instance, for high-value and loyal customers, personalized incentives or exclusive offers might be effective in maintaining their loyalty. Seasonal shoppers may respond well to targeted promotions during peak seasons.

Strategic Spenders could benefit from customized marketing campaigns highlighting complementary products or services. Meanwhile, customers identified as "at risk" might require re-engagement efforts, such as special offers to incentivize future purchases. The innovative implementation of the RFM model, coupled with ongoing customer behavior analysis, enables businesses to proactively address the needs and preferences of each customer category, ensuring a more effective and personalized recovery strategy. This approach fosters customer satisfaction, loyalty, and ultimately contributes to the overall success of the business.

## Conclusion

In conclusion, the significance of customer segmentation, particularly through the application of the RFM model, cannot be overstated. Recognizing the diverse nature of customer behaviors and preferences, effective segmentation serves as a cornerstone for targeted marketing strategies and enhanced customer engagement. The importance of innovative approaches in this context is paramount, considering the evolving dynamics of the market and the need for nuanced insights into customer interactions. In this study, we embarked on an exploration of innovative methods within the realm of customer segmentation. Our research aimed to push the boundaries of traditional approaches and introduce novel strategies for understanding and categorizing customer behaviors. Through meticulous investigation and analysis, we derived meaningful insights that contribute to the evolving landscape of customer-centric strategies.

The proposed RFM-based framework is not only effective for e-commerce segmentation but is also adaptable to other customer-centric industries such as banking, telecommunications, and insurance. In these domains, structured customer interaction data—such as account activity, payment frequency, or transaction value—can be mapped to Recency, Frequency, and Monetary dimensions, making the model suitable for applications like churn prediction, loyalty scoring, and personalized retention strategies. This cross-sector flexibility enhances the practical relevance of the framework and underscores its potential as a scalable decision-support tool.

Throughout the examination, it became evident that customers exhibited varying patterns, and the conventional methods of segmentation often fell short in capturing the intricacies of their behaviors. In response, our innovative approach allowed for a more accurate and detailed categorization of customers based on the RFM model. The results obtained underscored the prevalence of customer attrition to a significant extent. Moreover, the identification of customers within each segment was executed with precision, offering valuable insights that can inform strategic decision-making.

Through visualization, we sought to present these findings in a comprehensible and actionable manner. The visualization of results not only enhances the clarity of our discoveries but also provides businesses with a visual tool for interpreting and applying the insights derived from our study. In essence, our research not only underscores the importance of customer segmentation and innovative approaches but also contributes a valuable perspective to the ongoing discourse in the field.

This research embarked on a comprehensive exploration of customer segmentation using the

RFM model, delving into innovative approaches to understanding and responding to diverse customer behaviors. The research commenced with meticulous data collection, where a vast dataset, particularly from the United Kingdom, formed the foundation for subsequent analyses. The pre-processing phase ensured data cleanliness and relevance, paving the way for a robust implementation of the RFM model.

An innovative aspect of this research lies in the meticulous categorization of customers into six distinct segments, ranging from high-value patrons and loyal customers to strategic spenders and those identified as "at risk." The RFM scoring method, combined with strategic quartile divisions, contributed to a nuanced understanding of each customer's unique behavior. These insights were then visualized through informative charts, such as Figure 3, offering a dynamic representation of customer distribution across categories.

Furthermore, Figure 4 demonstrates the correlation between RFM indices and customer categories, enhancing our comprehension of the distinctive characteristics within each segment. Figure 6 provided an additional layer of insight by spotlighting mid-range RFM values across different customer categories. Exploring individual RFM indices in Figure 5 allowed for a deeper examination of the relationships between these indicators.

This study presents a transparent and interpretable RFM-based segmentation framework designed to support targeted customer retention and loyalty strategies. While the model was applied to a dataset from the UK e-commerce sector, its structure is not market-dependent and can be readily adapted to other industries such as banking, telecommunications, or insurance—where customer interaction data can similarly be mapped to Recency, Frequency, and Monetary dimensions. To further enhance its depth and precision, future research may integrate additional behavioral signals (e.g., clickstream data, browsing history, demographics) and apply advanced machine learning techniques such as survival analysis or ensemble learning. These extensions would not only improve predictive performance, particularly in Customer Lifetime Value (CLV) estimation, but also expand the framework's utility across diverse, data-rich environments.

## References

[1] M. Abirami and V. Pattabiraman, "Data mining approach for intelligent customer behavior analysis for a retail store," Proc. 3rd Int. Symp. Big Data Cloud Comput. Challenges (ISBCC), Springer, 2016, pp. 283–291.

[2] M. A. Ardakani, M. H. Karimi-Zarchi, and D. Shishebori, "Analyzing and forecasting of coronavirus time-series data: Performance comparison of machine learning and statistical models," Adv. Ind. Eng., vol. 58, no. 2, 2024, pp. 291–306.

[3] M. Frasquet, M. Ieva, and C. Ziliani, "Online channel adoption in supermarket retailing," J. Retail. Consum. Serv., vol. 59, 2021, Art. no. 102374.

[4] H. Gao, X. Li, and Y. Zhang, "A time-varying Cox model for dynamic churn risk prediction in subscription-based businesses," Expert Syst. Appl., vol. 230, 2024, Art. no. 120698.

[5] M. C. Gül and D. Şen, "Segmenting retail customers using RFM analysis and designing effective push notifications: A data-driven approach," Int. J. Retail Distrib. Manag., vol. 50, no. 3, 2022, pp. 321–339.

[6] S. Gupta, A. Patel, and S. Ray, "Transformer-based churn prediction in subscription-based businesses: A deep learning approach," J. Intell. Syst., vol. 9, no. 1, 2024, pp. 200–219.

[7] Ů. Hamidi, S. J. Sadjadi, A. B. Naeini, and S. R. Moosavi Tabatabaei, "An integrated production-marketing planning model with cubic production cost function and imperfect production process," J. Ind. Syst. Eng., vol. 10, Special Issue, 2017, pp. 91–108.

[8] M. Hernant and S. Rosengren, "The impact of adding an online channel to traditional retailing: A study of consumer behavior in a multichannel context," J. Retail. Consum. Serv., vol. 36, 2017, pp. 78–86.

[9] S. Khan, L. Zhu, X. Yu, Z. Zhang, M. A. Rahim, M. Khan, X. Du, and M. Guizani, "Accountable credential management system for vehicular communication," Veh. Commun., vol. 25, 2020, Art. no. 100279.

[10] A. Kumar, R. Bezawada, R. Rishika, R. Janakiraman, and P. K. Kannan, "From social to sale: The effects of firm-generated content in social media on customer behavior," J. Mark., vol. 80, no. 1, 2018, pp. 7–25.

[11] N. M. Larsen, V. Sigurdsson, and J. Breivik, "The use of observational technology to study in-store behavior: Consumer choice, video surveillance, and retail analytics," Behav. Analyst, vol. 40, 2017, pp. 343–371.

[12] C. Lee and K. Park, "Comparing traditional and ensemble learning methods for churn prediction in the telecom industry," J. Bus. Anal., vol. 6, no. 2, 2024, pp. 90–105.
[13] K. Melis, K. Campo, E. Breugelmans, and L. Lamey, "The impact of online versus offline channel additions on customer behavior," Int. J. Res. Mark., vol. 33, no. 3, 2016, pp. 725–735.
[14] M. Mohammadi, M. R. Rasouli, and M. S. Pishvaee, "Data driven approaches for customer centric and service dominant value propositions: A systematic literature review," J. Ind. Syst. Eng., vol. 14, no. 1, 2021, pp. 51–75.
[15] E. Montaguti, S. Kuester, and V. R. Rao, "Advertising investments in a recessionary environment: Enhancing customer loyalty via multichannel strategies," J. Mark., vol. 80, no. 3, 2016, pp. 89–106.
[16] E. Park, Y. Jang, J. Kim, N. J. Jeong, K. Bae, and A. P. Del Pobil, "Determinants of customer satisfaction with airline services: An analysis of customer feedback big data," J. Retail. Consum. Serv., vol. 51, 2019, pp. 186–190.
[17] K. Pauwels and S. A. Neslin, "Building with bricks and mortar: The revenue impact of opening physical stores in a multichannel environment," J. Retail., vol. 91, no. 2, 2015, pp. 182–197.
[18] M. A. Rahim, M. Mushafiq, S. Khan, and Z. A. Arain, "RFM-based repurchase behavior for customer classification and segmentation," J. Retail. Consum. Serv., vol. 61, 2021, Art. no. 102566.
[19] J. R. Saura, P. R. Palos-Sanchez, and M. B. Correia, "Using machine learning and sentiment analysis to improve customer segmentation and product strategies," Technol. Forecast. Soc. Change, vol. 190, 2023, Art. no. 122331.
[20] S. E. Schaeffer and S. V. R. Sanchez, "Forecasting client retention—A machine-learning approach," J. Retail. Consum. Serv., vol. 52, 2020, Art. no. 101918.
[21] T. Wang, "Enhancing digital marketing through customer segmentation using reinforcement learning and purchase frequency metrics," J. Mark. Anal., vol. 13, no. 2, 2025, pp. 109–127.
[22] H. Y. Wong, Y. Tan, and T. Teo, "Integrating clustering techniques with behavioral metrics for e-commerce customer segmentation," J. Bus. Res., vol. 165, 2024, Art. no. 112984.
[23] Y. Xu, L. Zhang, and H. Liu, "Deep clustering-based RFM segmentation for e-commerce personalization," Electron. Commer. Res. Appl., vol. 58, 2023, Art. no. 101152.
[24] D. Zammit and C. Zerafa, "A simplified and numerically stable approach to the BG/NBD churn prediction model," arXiv Prepr., arXiv:2502.12912, 2025.
[25] M. Ziari, "A data-driven pricing model for distribution systems considering competition," Adv. Ind. Eng., vol. 58, no. 1, 2024, pp. 179–195.