Customer Segmentation for Online Retail Using Clustering Algorithms

Qazi Noman
University of South Wales
South Wales
30122401@students.southwales.ac.uk

***Abstract*—This study set out to build a practical customer segmentation approach for online retailers, tapping into transactional data and clustering techniques. We worked with a rich dataset from a UK-based online gift retailer, covering over 541,000 transactions by 4,372 customers between December 2010 and December 2011. Instead of relying on - [Truncated by character limit, please expand to read more] traditional demographic-based segmentation, we focused on customer behavior, using RFM (Recency, Frequency, Monetary value) analysis alongside metrics like product diversity and purchase timing patterns to capture how customers shop.**

We tested three clustering methods—K-means, Hierarchical, and DBSCAN—to uncover natural customer groups. K-means clustering, producing five segments, struck the best balance between statistical rigor (silhouette score: 0.306) and practical business value. The segments we identified were Champions (31.7 For each group, we crafted tailored marketing strategies: retention plans for Champions to keep them engaged, deeper engagement tactics for Loyal Customers, and re-engagement campaigns to win back Lost Customers. We also mapped out implementation steps, covering technical needs, system integration, and resource planning to ensure these strategies deliver strong returns on marketing investment.

This research advances retail analytics by offering a clear, behavior-driven segmentation framework that ties clustering methods to actionable marketing plans. It empowers online retailers to shift from generic, one-size-fits-all campaigns to personalized strategies that boost customer satisfaction and drive business growth.

***Index Terms*—customer segmentation, RFM analysis, clustering algorithms, behavioral analytics, online retail, marketing strategies, K-means clustering.**

I. Introduction

A. Background and Motivation

The retail industry has transformed dramatically over the past two decades, propelled by the rapid growth of online shopping and the shift toward digital customer interactions. Global e-commerce sales soared from $1.3 trillion in 2014 to $4.9 trillion in 2021, with projections estimating sales will surpass 7.4 trillion $ by 2025 (Statista, 2022). This remarkable change has redefined how retailers connect with customers, collect information, and design marketing strategies. Traditional physical stores had limited access to customer data, often restricted to basic details captured during sales. In contrast, online retailers now have access to a wealth of digital insights, tracking behaviors like browsing habits, abandoned carts, purchase frequency, and even post-purchase interactions. This data-rich environment has raised customer expectations significantly. Shoppers today want retailers to understand their unique preferences and deliver personalized experiences. A 2021 McKinsey study revealed that 71% of consumers expect tailored interactions, and 76% feel frustrated when these expectations aren't met (McKinsey, 2021). As a result, effective customer segmentation—grouping customers based on shared traits—has become essential for online retailers, evolving from a competitive edge to a core business necessity. Customer segmentation isn't new; it began in the 1950s with a focus on demographics like age, gender, income, or location. While these categories offered some insights, they often failed to predict how customers would shop. Two people with similar demographic profiles might have completely different buying habits, limiting the usefulness of demographic-based groups for targeted marketing. The digital era has shifted segmentation toward behavior, focusing on what customers do—such as how often they buy, what they browse, or which products they prefer. This approach aligns more closely with the true goal of segmentation: identifying groups that respond similarly to marketing campaigns or product offerings. The flood of transactional data from online retail platforms provides a unique opportunity for advanced segmentation. Every click, purchase, return, or review generates valuable insights that traditional methods couldn't capture. Retailers use machine learning tools with clustering algorithms to discover patterns in their data which transforms their strategies from guesses into data-based decisions. RFM analysis stands as a powerful method which evaluates three essential customer metrics: Recency (how recently a customer shopped), Frequency (how often they purchase) and Monetary value (how much they spend). The combination of behavioral details such as product variety and transaction timing with sophisticated algorithms enables RFM to establish meaningful customer groups that drive effective marketing strategies. The retail industry undergoes fundamental changes because data and technology now influence business operations. Retailers who implement these tools will deliver personalized experiences to customers which build loyalty and generates sales. The complete utilization of this data needs to overcome major obstacles which this research tackles by developing a systematic behavioral segmentation methodology.

B. Problem Statement

While the promise of data-driven segmentation is clear, putting it into practice poses real challenges for online retailers. Moving from demographic to behavioral segmentation introduces complexity that many businesses struggle to manage. Demographic data offers straightforward, stable categories—age or location don't change rapidly. Behavioral data, on the other hand, is dynamic and multifaceted, capturing customer actions across websites, apps, devices, and touchpoints. This makes it harder to create clear, actionable groups. The sheer volume of data in online retail adds both opportunity and difficulty. A medium-sized retailer might process thousands of transactions daily, each packed with details about products, purchase timing, payment methods, and customer interactions. Without the right analytical tools, this wealth of data becomes a burden rather than an asset. Many retailers collect vast amounts of information but lack the expertise or framework to turn it into practical segmentation strategies. A 2022 Forrester survey highlighted this gap, noting that 74% of companies aim to be data-driven, yet only 29% successfully connect analytics to actionable outcomes (Forrester, 2022). This challenge is especially evident in customer segmentation, where businesses may have advanced data systems but still rely on basic methods that fail to capture the full complexity of customer behavior. Choosing the right clustering algorithms is another hurdle. Different methods—like K-means, which groups customers around central points; Hierarchical clustering, which builds groups step by step; or DBSCAN, which identifies dense clusters—produce different results based on their underlying assumptions. Without explicit evaluation criteria, retailers can develop segments that are statistically valid but meaningless to their business requirements. Interpreting technical findings into useful segments is no less difficult. Clustering algorithms yield measures such as silhouette scores or inertia, which confirm the math but don't necessarily explain what the segments imply for marketing. A statistically ideal clustering may be worthless if it doesn't support a retailer's marketing capabilities or objectives. Filling this gap between data science and marketing strategy is still a major obstacle. Lastly, it is inherently challenging to measure the effect of segmentation. Unlike dedicated marketing campaigns with clean metrics such as click-through rates, the pay-perks of enhanced segmentation like higher retention or increased customer lifetime value emanate indirectly. This makes it difficult for retailers to make an economic case for investing in more sophisticated segmentation methods, which have encouraged many to cling to less complicated, less productive approaches despite their own shortcomings. Resolution of these challenges depends on having a clear, structured framework which not only identifies groups of customers but also renders them actionable in real-world retail.

C. Research Objectives

The current research seeks to address these concerns by creating a sound customer segmentation framework in online retailing, based on transaction data to reveal behaviourally different customer groups. The main aim is to determine these groups and assess how well various clustering techniques uncover meaningful patterns. This aim is underpinned by several specific objectives that inform the research. Firstly, we aim to determine the usefulness of RFM analysis for online retail segmentation. Although RFM has been shown to be useful in other retail environments, online shopping is different, with features such as quick purchase cycles and diverse customer expectations. We'll evaluate how well Recency, Frequency, and Monetary metrics—both individually and together—identify distinct customer groups. To enhance RFM, we'll also explore additional behavioural metrics, such as the range of products purchased, the timing of transactions, and the composition of shopping carts, to create a more comprehensive view of customer behaviour. Second, we'll compare three clustering methods: K-means, Hierarchical, and DBSCAN. They all cluster differently—K-means is based on central points, Hierarchical constructs clusters piecewise, and DBSCAN seeks dense groups. We will test their performance against technical measures such as silhouette scores and the Calinski-Harabasz index, as well as their pragmatic utility for retail. This analysis will assist us in deciding which method best weighs accuracy against business relevance. Third, we intend to create elaborate profiles for each customer group, going beyond numbers to describe what makes them unique. These profiles will highlight their behaviours, their value to the business, their potential for growth, and any risks, such as churn. By understanding these traits, retailers can develop marketing strategies that resonate with each group's specific needs and habits. Finally, we'll create tailored marketing recommendations for each segment, emphasizing engagement, retention, and growth. These will involve recommendations regarding frequency of communication, most desired channels (e.g., email or social media), discounting strategies, and product recommendations tied to the behaviour of each group. We will also offer practical implementation guidance, including technical requirements, resource requirements, and how to measure success, so that retailers are able to effectively implement the framework.

D. Significance of Study

This work makes valuable contributions to academic research and real-world retail practice. Academically, it fills a knowledge gap in understanding how classical segmentation techniques work in the multifaceted, behaviour-led world of e-commerce. Through experimenting with several clustering methods within this setting, it presents novel insights and methodology for researchers and data scientists who practice retail analytics. Comparison of clustering techniques enlightens their advantages and limitations when implemented to actual retail data, which in most cases are not evenly structured and have their own set of problems. Though clustering is computer science's best-studied field, retail demands careful management of data preprocessing, feature subset selection, and assessment. Through addressing these problems and developing methods for balancing technical accuracy with business utility, this research contributes to the practice of applied data science in retailing. For retailers, this research gives a detailed, step-by-step process for developing sophisticated, behaviour-based segments. It spans from data preparation to algorithm selection to profiling segments, so that it can be used by companies at any stage of data competence. The marketing suggestions deal with a general shortcoming of segmentation studies, which tend to end identifying groups without specifying how to respond to them. By connecting segment attributes with precise strategies—like communication approaches, promotional promotions, and resource allocation plans, this research supports the ability of retailers to convert analytics into tangible outcomes. The study also contributes to the wider retail environment, including technology players, marketing firms, and pedagogical offerings. By establishing benchmarks for segmentation and measurement, it aids in creating improved tools and services that are specific to retail requirements. It also guides retail education, preparing future leaders to acquire the data-driven competencies necessary in the modern competitive environment. Above all, improved segmentation serves customers by allowing more pertinent, personalized shopping. When retailers know what various customers desire, they can provide products, messages, and interactions that are custom-seeming and significant. This results in a win-win: consumers have improved experience, and retailers gain loyalty, enhance performance, and prosper in the fast-paced online environment.

II. Literature Review

A. Customer Segmentation Techniques

The idea of market segmentation took shape in 1956 when Wendell Smith proposed that large, diverse markets could be broken into smaller, more similar groups based on customers' unique preferences. This was a bold shift from the one-size-fits-all marketing of time, laying the groundwork for modern marketing strategies. Smith's insight—that customers have different needs, tastes, and behaviors—pushed businesses to tailor their approaches to specific groups rather than treating everyone the same. Segmentation has evolved through distinct stages, starting with broad categories like geography and demographics. Early efforts focused on straightforward traits such as where people lived, their age, gender, income, or job. These were easy to measure and apply, but they often missed the mark on explaining why people bought what they did. As Wind (1978) pointed out, demographics give a starting point for grouping customers, but they rarely reveal the deeper reasons behind purchasing choices. By the 1970s and 1980s, researchers began exploring psychographic segmentation, which dug into customers' lifestyles, personalities, and values. The VALS (Values, Attitudes, and Lifestyles) framework became a popular tool, sorting people into categories like "Achievers," "Experiencers," or "Believers" based on their mindsets and life goals. This approach offered richer insights into what motivated customers, but it was tricky to measure and apply in practice, often requiring complex surveys or interviews. The 1990s brought behavioral segmentation to the forefront, fueled by better data from point-of-sale systems and loyalty programs. Instead of focusing on who customers were, this method looked at what they did, how often they bought, whether they stuck with certain brands, how much they used a product, or what benefits they sought. Kotler and Keller (2012) emphasized that behavioral data is the best foundation for building market segments because it ties directly to actions that businesses care about, like purchases and engagement. The rise of online shopping has supercharged behavioral segmentation. E-commerce platforms track every step of a customer's journey, from the moment they land on a website to browsing, comparing products, making a purchase, and even leaving reviews. This wealth of data allows businesses to move beyond static snapshots of customers and instead analyze their behavior over time, capturing shifts and patterns that older methods couldn't detect. Today's segmentation often blends multiple angles, creating hybrid models that combine demographics, psychographics, and behaviors. These approaches recognize that no single lens fully captures why customers act the way they do. Tsiptsis and Chronopoulos (2011) argue that the best segmentation mixes these dimensions to create groups that are distinct, easy to identify, reachable, large enough to matter, and useful for marketing. This multidimensional view helps businesses craft strategies that resonate with specific customer needs. In online retail, segmentation has shifted from occasional, research-driven projects to ongoing, data-driven processes. Traditional stores relied on surveys, focus groups, or limited sales data to understand customers. E-commerce, on the other hand, automatically collects detailed behavioral data without needing direct customer input. This has made segmentation a core part of daily operations, powered by algorithms that continuously refine customer groups. Despite these advances, there's still a gap between academic research and real-world retail. Scholars have developed complex segmentation methods, but many businesses stick to simpler, rule-based approaches due to limited skills, outdated systems, or resistance to change. Another challenge is turning statistical results into marketing plans that work. Academic studies often focus on math, while retailers need clear, practical strategies. The cutting edge of segmentation lies in using machine learning to uncover hidden customer patterns. Unlike rigid, predefined rules, these methods let algorithms find natural groupings in complex behavioral data. Combining these tools with marketing expertise creates a powerful approach, blending the precision of data science with the practical know-how of retail. This balance is key to building segments that are both statistically sound and useful for business.

B. RFM Analysis

Recency, Frequency, Monetary (RFM) analysis stands out as a practical and widely used tool for behavioral segmentation in retail. Introduced by Hughes (1994), RFM looks at three core aspects of customer behavior: how recently someone made a purchase (Recency), how often they buy (Frequency), and how much they spend (Monetary value). Its staying power comes from its direct link to key business goals—keeping customers engaged, building loyalty, and driving revenue. RFM's strength lies in its grounding in real-world customer patterns. Recency is a strong predictor of future purchases because people who bought recently are more likely to buy again, reflecting habits and brand familiarity. Fader et al. (2005) found that recency often outperforms more complex metrics in predicting future buying, making it a simple yet powerful tool. This aligns with psychological ideas about how recent actions reinforce ongoing behavior. Frequency shows how engaged or loyal a customer is. Those who buy often tend to have a stronger connection to the brand and are less likely to switch to competitors. Reichheld (2001) showed that boosting customer retention (tied closely to frequency) by just 5% can lift profits by 25% to 95%, underscoring why this metric matters. Frequency helps businesses identify their most dependable customers and focus efforts on keeping them. Monetary value measures how much a customer contributes financially, hinting at their budget, spending habits, or interest in a product category. High-spending customers might have more disposable income, greater decision-making power, or a deeper connection to the products. Together, these three metrics paint a detailed picture of a customer's value, balancing their past contributions with their potential for future business. Traditionally, RFM analysis splits customers into groups (often quintiles) for each metric, creating a three-digit score. A score of 555, for instance, means a customer is in the top group for recency, frequency, and monetary value—a highly valuable customer. With quintiles, this creates 125 possible combinations, which businesses typ- ically simplify into a few key segments for marketing. This method is simple and simple to use, which makes it popular in all industries. But classic RFM has its limits, particularly in sophisticated retail environments. Quintile cutoffs can be arbitrary, clustering customers with slight differences while dividing those with comparable behaviours. It also presumes recency, frequency, and monetary value are equally important, which isn't always the case. For instance, in luxury retail, spending may be more important than frequency, whereas at grocery stores, recent buys may be more significant. Contemporary RFM techniques eliminate these problems. Weighted RFM gives various importance to every measure depending upon the business environment, making sure the model behaves as it should in the real world. Time-adjusted RFM considers how long the customer has remained active, without biasing new customers who did not have a chance to accumulate purchases. These refinements make RFM more equitable and more appropriate. Combining RFM with clustering algorithms is a game-changer. Rather than rigid quintile lines, clustering identifies natural groups of customers based on their RFM patterns, preventing artificial bifurcations and revealing segments that could cut across several quintile levels. This methodology allows businesses to uncover more subtle groups of customers that standard RFM could overlook. Researchers have further enriched RFM by including new measurements. Chang and Tsay (2004) developed LRFM, which incorporates the duration of the customer's relationship with the brand to provide a more adequate measure of loyalty. Purchase timing (e.g., weekdays or weekends) has been added by some to identify patterns that inform when to send promotions. These additions make RFM more dynamic and customisable to retail requirements. In online shopping, RFM excels because the digital environment offers boundless data—every transaction, click, or cart abandonment is monitored. But online shopping activities, such as changing devices or browsing without purchasing, need adjustments to conventional RFM. Peker et al. (2017) demonstrated that incorporating digital measures, such as website visits or browsing time, enhances RAM's performance in e-commerce, producing more acute, actionable segments. Nevertheless, RFM is not flawless. It looks at past behaviors, overlooking attitudes or preferences that may influence future buying. It also handles purchases in isolation, without regard to details such as product classes or purchasing sequences that might uncover richer knowledge. Such deficiencies have prompted researchers to look for ways to merge RAM's simplicity with more sophisticated tools, remaining topically relevant in the modern data-driven retail environment.

C. *Clustering Algorithms in Retail*

Clustering algorithms have become essential for retail segmentation, helping businesses uncover natural customer groups from complex, multidimensional data. Unlike rule-based methods that rely on preset criteria, clustering uses statistical patterns to group customers with similar behaviors. This approach reveals segments that might not be obvious through traditional marketing analysis, offering a more precise view of customer diversity. K-means clustering, introduced by MacQueen (1967), is the go-to algorithm for retail due to its speed and clarity. It groups customers into a set number of clusters (K) by minimizing the differences within each group, creating clusters where customers share similar traits, like RFM scores. K-means works well with the typical data patterns in retail and produces clear, interpretable results. Hosseini et al. (2010) used K-means with RFM data for a retail chain and saw better response rates to promotions compared to demographic-based groups. Chen et al. (2012) applied it to an online retailer, improving campaign performance by 25% over unsegmented strategies. But K-means has drawbacks. It requires choosing the number of clusters upfront, which can miss the data's natural structure. It also assumes clusters are roughly round and similar in size, which doesn't always match retail data, where customer groups can vary widely or have irregular shapes. Outliers, like unusually high spenders, can also skew results. Hierarchical clustering offers a different approach, building clusters either from the ground up (grouping individual customers) or top down (splitting a single large group). The bottom-up method, especially with Ward's technique, is popular in retail because it minimizes differences within clusters as it builds them. Its biggest strength is the dendrogram—a visual tree showing how clusters connect and how distinct they are. Ballestar et al. (2018) showed that hierarchical clustering can uncover layered segment structures, letting retailers choose how detailed their segmentation needs to be based on marketing goals or resources. DBSCAN (Density-Based Spatial Clustering of Applications with Noise), developed by Ester et al. (1996), excels at finding clusters of any shape and spotting outliers. It groups customers in dense areas of data, separated by sparser regions, and flags those who don't fit any group as outliers. This is especially useful in retail for finding niche segments or unusual customers. Christy et al. (2018) used DBSCAN as a specialty retailer and found valuable micro-segments that other methods overlooked, as well as high-value or risky customers needing special Attention. Model-based clustering, like Gaussian Mixture Models (GMM), assumes customers come from a mix of statistical distributions. Unlike K-means or DBSCAN, GMM lets customers belong to multiple segments with varying probabilities, reflecting the fluid nature of shopping behavior. This flexibility makes it a good fit for retail, where customers often don't fit neatly into one box. Choosing the right algorithm depends on the retail context. K-means works well for simpler datasets, while DBSCAN handles complex, high-dimensional data better. Large retailers with millions of customers need algorithms that scale efficiently, which often favor K-means. Marketing teams, who may not be stats experts, often prefer simpler methods for easier interpretation, even if they sacrifice some precision. Recent trends in retail clustering focus on blending algorithms with business knowledge. Semi-supervised cluster-ing uses some customer labels or business rules to steer the process toward practical results. Ensemble clustering combines multiple algorithms to create stronger, more reliable segments, balancing the strengths and weaknesses of each. These hybrid approaches help ensure segments are both data-driven and business-relevant.

D. *Evaluation Metrics for Clustering*

Evaluating clustering results is tricky because, unlike tasks with clear right-or-wrong answers, clustering always produces groups, whether they're meaningful or not. In retail, the real test of segmentation is whether it drives business results, but statistical metrics are still crucial for assessing how well clusters reflect customer patterns. The challenge is balancing these technical measures with practical marketing needs. Internal metrics look at the data and clusters themselves, without needing outside reference points. The silhouette coefficient, introduced by Rousseeuw (1987), measures how similar customers are within their cluster (cohesion) compared to other clusters (separation). Scores range from -1 to 1, with higher values showing tighter, more distinct groups. This is widely used in retail to check if segments are well-defined. The Davies-Bouldin index compares each cluster to its closest neighbor, with lower scores indicating more separated groups. This helps retailers spot when segments look different but behave similarly, avoiding redundant marketing efforts. The Calinski-Harabasz index measures how spread-out clusters are compared to how tight they are internally. Higher scores mean clearer, more compact segments, which marketers prefer for targeting. It works well for K-means but less so for DBSCAN, which creates non-round clusters. These statistical metrics are helpful but don't always tell the whole story. Mathematically perfect clustering might group customers in ways that don't make sense for marketing, while a less perfect one might reveal practical insights. To bridge this gap, retailers use business-focused criteria: actionability (do segments suggest clear strategies?), substantiality (are segments big enough to target?), stability (do segments hold up over time?), and distinctiveness (do segments respond differently to marketing?). Cross-validation methods add rigor. Holdout validation splits customers into two groups, clusters one set, and tests how well, the other fits those clusters. Temporal validation checks if segments built from past data still make sense later, ensuring they're not just temporary patterns. These approaches help retailers focus on lasting, valuable segments. Visualization tools offer a hands-on way to evaluate clusters. Principal Component Analysis (PCA) shrinks complex data into two or three dimensions for plotting, letting marketers see if clusters are well-separated. t-SNE, another visualization method, keeps local patterns intact, showing whether customers form clear groups or blend together. These visuals help non-technical teams grasp the results. Deciding how many clusters to use is another key step. The elbow method plots the number of clusters against the within-cluster variation, looking for a point where adding more clusters doesn't add much value. The gap statistic compares the data's clustering to a random baseline, offering a more formal way to pick the right number of segments. External validation, comparing clusters to known customer labels, is rare in retail since true labels are often missing. But in research, metrics like the Adjusted Rand Index or Normalized Mutual Information can test how well clusters match expected groups in controlled settings. Ultimately, good clustering evaluation combines multiple angles. Statistical metrics ensure clusters are well-formed, business criteria check their marketing value, and time-based validation confirms they're reliable. This layered approach helps retailers avoid segments that look good on paper but fall short in practice, or ones that feel intuitive but lack data backing.

III. Methodology

A. Dataset Description

For this research, I selected the" Online Retail" dataset, a rich and detailed collection of transactional records from a UK-based online retailer specializing in unique, all-occasion gifts. This dataset, generously contributed by Dr. Daqing Chen from London South Bank University to the UCI Machine Learning Repository, stands out as an exceptional resource for customer segmentation studies. Its real-world context, comprehensive scope, and manageable size make it perfect for in-depth analysis without demanding the kind of computational infrastructure required for massive datasets. Working with this dataset felt like stepping into the messy, fascinating reality of retail, where customer behaviors unfold in unpredictable yet revealing ways. The dataset captures every transaction from December 1, 2010, to December 9, 2011—a full year plus a bit extra, which I found ideal for studying customer patterns. This timeframe encompasses multiple purchasing cycles, seasonal peaks like Christmas, and enough longitudinal data to reveal meaningful trends in how customers shop. With 541,909 transaction records from customers in 38 countries, the dataset offers a broad perspective, though it's heavily weighted toward the UK, which accounts for roughly 91% of sales. This UK dominance shaped my approach, as it suggested a homogeneous market with some international diversity to spice things up. Each transaction record includes eight variables that together provide a clear snapshot of the purchase event: 1. Invoice No: A 6-digit code uniquely identifying each transaction. If it starts with a 'C', it signals a cancellation, which I had to account for carefully. 2. Stock Code: A 5- digit code assigned to each product, helping track what was bought. 3. Description: The product's name as it appears in the catalog, adding a human touch to the data. 4. Quantity: The number of units purchased in a single transaction. 5. Invoice Date: The precise date and time of the transaction, critical for temporal analysis. 6. Unit Price: The price per unit in pounds sterling (£), essential for monetary calculations. 7. CustomerID: A 5-digit number uniquely tied to each customer, the backbone of segmentation. 8. Country: The customer's country of residence, offering geographic context. What makes this dataset particularly compelling for segmentation is its quirks. About 25% of transactions lack a CustomerID, which I suspect reflects anonymous purchases or technical glitches in data collection. This missing data posed a challenge but also mirrored the messy reality of retail systems. Additionally, the purchasing patterns are highly skewed—a small group of customers generates a disproportionate share of transactions and revenue. This skewness, typical in retail, is exactly what makes segmentation so valuable, as it highlights the need to identify and prioritize high-value customers. However, the dataset has limitations that shaped my methodology. There's no demo- graphic information, so I couldn't explore how age, gender, or income might influence buying behavior. It also lacks data on browsing habits, product views, or cart abandonments, which would have enriched the customer journey analysis. Despite these gaps, the transactional data is robust enough to support segmentation based on purchasing patterns, offering a solid foundation for my research objectives. The dataset's real-world nature, complete with missing values, outliers, and natural variations, made it an authentic testing ground for clustering algorithms. The year-long span allowed me to dive into recency analysis with confidence, while the detailed trans- action records opened the door to creative feature engineering beyond the standard RFM (Recency, Frequency, Monetary) framework. I saw this dataset as a puzzle—imperfect but full of potential to reveal actionable insights about customer behavior.

B. *Data Preprocessing Steps*

Getting the raw transaction data ready for clustering was no small task. Retail datasets like this one are notorious for quality issues—missing values, negative quantities, out- liers, and formatting inconsistencies—and I needed to address each systematically to ensure the analysis would hold up. My preprocessing workflow was designed to clean the data thoroughly while preserving its integrity, setting the stage for reliable clustering results. The first hurdle was missing CustomerIDs, which affected 24.93% of transactions—roughly 135,080 records. Since customer segmentation relies on group- ing transactions by individual customers, this was a significant issue. I explored options like imputing IDs based on patterns in IP addresses or payment methods, but these felt speculative and risky. Ultimately, I decided to exclude transactions without CustomerIDs. This reduced the dataset to 406,829 records, a tough but necessary choice to ensure every transaction could be tied to a specific customer. I felt this preserved the authenticity of the segmentation process, even if it meant losing some data. Next, I tackled cancelled orders, identifiable by negative quantities or invoice numbers starting with 'C'. These represent returns or cancellations, not actual purchases, and including them would have distorted monetary and frequency metrics. For example, a cancelled order could artificially lower a customer's total spending, skewing their value. I removed 8,287 such records, refining the dataset to focus solely on completed purchases. This step felt like clearing away noise to reveal the true signal of customer behavior. Data cleaning went beyond these big issues to address smaller but equally important problems I uncovered during initial exploration. Duplicate transactions, likely from system errors, showed up as identical InvoiceNo, CustomerID, StockCode, and Quantity values within the same minute. I carefully removed these to avoid inflating purchase counts. Transactions with unusually high quantities—beyond three standard deviations from the mean—were another red flag. I reviewed these individually, correcting or dropping those that seemed like data entry errors rather than legitimate bulk purchases. This process required judgment, as some high- quantity orders were clearly real and valuable for segmen- tation. I also found transactions with zero or negative unit prices, which would throw off monetary calculations. These, likely promotional items, samples, or errors, were removed. Transactions with unit prices exceeding £10,000 were scrutinized to confirm they were genuinely high value purchases, not mistakes, as such outliers could heavily influence clustering results if erroneous. This step was particularly delicate, as I wanted to respect the dataset's natural variability while ensuring accuracy. Outliers posed a unique challenge given the dataset's skewness. In retail, a small number of high-value customers naturally stand out, and standard statistical methods assuming normal distributions would wrongly flag them as errors. Instead, I adopted a domain-specific approach, focusing on anomalous patterns rather than raw statistical deviations. Customers with sudden, extreme changes in behavior—like a one-time massive purchase—were flagged for review, while consistently high-value customers were kept, as they represent a critical segment. This felt like balancing statistical rigor with retail reality, ensuring I didn't lose valuable insights. Temporal consistency was another priority. I converted all InvoiceDate values to a standard datetime format and extracted features like day of week, month, quarter, and time of day to support temporal analysis. The dataset spanned December 1, 2010, to December 9, 2011, and I set December 10, 2011, as the reference date for recency calculations to ensure uniformity across customers. This step was straightforward but essential for aligning the data for analysis. The final preprocessing task was aggregating transaction-level data into customer-level data for clustering. I combined all transactions for each customer into a single record summarizing their purchasing history and derived metrics. This process yielded a dataset of 4,372 unique customers with complete behavioral profiles, ready for segmentation. Looking at this clean, customer-level dataset felt like reaching a milestone—after all the wrangling, I had a solid foundation for the next steps.

C. *Feature Engineering (RFM and Additional Features)*

Feature engineering was where the dataset came to life, transforming raw transactions into meaningful customer attributes that capture distinct buying behaviors. I started with the tried-and-true RFM framework but didn't stop there, adding custom features to uncover deeper insights into how customers shop. This process was both technical and creative, as I sought to translate numbers into a story about customer habits. Recency measured the days since a customer's last purchase, using December 10, 2011, as the reference point. This metric reflects how recently a customer engaged with the retailer, a strong indicator of their likelihood to buy again. I calculated it by subtracting the customer's latest InvoiceDate from the reference date, producing a continuous variable where lower values mean more recent activity. Recency ranged from 1 to 373 days, with a median of 50 and a right-skewed distribution, suggesting most customers had shopped recently. This skewness made sense—active customers tend to cluster toward recent purchases, while others drift away over time. Frequency counted the number of unique invoices per customer, focusing on shopping occasions rather than total items bought. This distinction was critical: a customer buying ten items in one transaction gets a frequency of one, while ten single-item purchases over time score ten. This approach highlights true frequent shoppers versus occasional bulk buyers, which matters for targeting strategies. Frequency ranged from 1 to 208, with a median of 2 and significant positive skew, reflecting the retail norm where a few customers shop often while most are sporadic. Monetary value summed the total spend (Quantity × Unit Price) per customer, capturing their overall contribution rather than average trans- action size. This aligns with the goal of identifying high- value customers who drive revenue. Monetary values ranged from £3.75 to £280,206.02, with a median of £673.10 and extreme skewness. This wide range underscored the need for logarithmic transformation to prevent high-value outliers from dominating clustering results. Seeing such a massive range was a reminder of how diverse customer behaviors can be in retail. Beyond RFM, I engineered additional features to capture more nuanced behaviors, drawing on my understanding of retail dynamics: Product diversity was measured in two ways. UniqueProducts counted distinct products purchased, ranging from 1 to 1,603, showing how broadly customers shopped. ProductDiversity, the ratio of unique products to total items, ranged from 0.001 to 1.0, distinguishing customers who stick to Favorites from those exploring the catalog. These metrics felt like a window into customer curiosity and loyalty. Purchase timing features captured when customers shop. WeekendProportion calculated the percentage of purchases on weekends versus weekdays, revealing preferences that could inform marketing campaigns. DayOfWeekVariation used the coefficient of variation to measure consistency in shopping days, identifying whether customers had predictable routines. SeasonalityScore assessed purchase distribution across quar- ters, separating seasonal shoppers from year-round regulars. These features added a temporal dimension that RFM alone couldn't capture. Order size metrics included ItemsPerTrans- action (average items per order), QuantityStdDev (variability in order size), and MaxQuantity (largest single order). These distinguished steady, small-order customers from those with sporadic bulk purchases, which has implications for inventory and promotions. I found these metrics particularly insightful for understanding shopping styles. Average order value (AOV) calculated the mean transaction value per invoice, complementing total monetary value. Two customers with identical total spend might differ—one making frequent small purchases, another rare large one—a distinction that shapes marketing approaches. AOV helped me see the granularity of customer value. To prepare for clustering, I addressed skewness in features like Monetary, Frequency, and order size metrics with logarithmic transformations, reducing the impact of extreme values. All features were then standardized using z-scores to ensure equal weighting in clustering algorithms, preventing larger-scale metrics from dominating distance calculations. This step felt like leveling the playing field, ensuring every feature had a fair shot at influencing the results.

D. Clustering Algorithms Implementation

To uncover natural customer segments, I implemented three clustering algorithms—K-means, Hierarchical, and DB- SCAN—each with a different approach to grouping cus- tomers. Testing multiple methods allowed me to compare their strengths and uncover diverse patterns in the data, ensuring a robust segmentation process. This stage was both exciting and challenging, as I navigated the nuances of each algorithm to find the best fit for retail segmentation.

K-means clustering was implemented using scikit-learn's KMeans class, with 'k-means++' initialization to select centroids strategically and improve convergence. I tested cluster counts (k) from 2 to 10 to determine the optimal number of segments, running each k with 10 different initializations (init=10) and up to 300 iterations to minimize the impact of random starting points. Euclidean distance was used, aligning with the normalized features from preprocessing. Tuning k involved the elbow method—plotting within-cluster sum of squares (WCSS) against k to spot diminishing returns—and silhouette scores to quantify clustering quality. These comple- mentary approaches provided a clear path to selecting the best k, balancing complexity and insight.

Hierarchical clustering used scipy's linkage function with Ward's method, which minimizes within-cluster variance, making it well-suited for segmentation. The algorithm starts with each customer as a single cluster, iteratively merging the closest pairs to build a dendrogram that visualizes cluster relationships at different levels. I tested four linkage methods—Ward's, complete, average, and single linkage—evaluating their dendrograms for clear, interpretable segments. Ward's method stood out for producing distinct, meaningful groups, while single linkage led to chaining, blurring boundaries. The dendrogram felt like a map of customer relationships, guiding my segmentation choices. DBSCAN, implemented via scikit-learn's DBSCAN class, took a density-based approach, identifying clusters based on data density rather than a preset number. This required tuning two parameters: epsilon and min-samples (the minimum points for a dense region). I used a k-distance graph to find the elbow point for epsilon, testing values (0.5, 0.75, 1.0, 1.25, 1.5) alongside min-samples (5, 10, 15) to balance meaningful clusters with minimal noise. DBSCAN's ability to flag outliers as noise was appealing, but tuning it felt like a delicate dance to avoid over- or under-clustering. Several challenges emerged during implementation. The 23-feature dataset risked the "curse of dimensionality," where distance measures lose meaning in high-dimensional spaces. To address this, I applied Principal Component Analysis (PCA) to reduce dimensions while preserving 80% of variance, improving both efficiency and interpretability. A high-value customer with an exceptionally large monetary value posed another issue. Rather than removing this legitimate outlier, I used log transformations and robust scaling to mitigate its impact, preserving valuable data. These solutions required balancing technical precision with retail context. The analysis was conducted in Python 3.8 within a Jupyter Notebook, using pandas for data handling, scikit-learn for machine learning, scipy for hierarchical clustering, matplotlib and seaborn for visualizations, and yellowbrick for clustering-specific plots. This environment allowed iterative exploration and visual checks, making the process both rigorous and intuitive.

E. Validation Approach

Validating the clustering results was critical to ensure the segments were statistically sound and practically useful. I designed a comprehensive validation strategy combining technical metrics, visualizations, and business-oriented evaluation, aiming to bridge data science with marketing applications. This stage felt like putting the segments under a microscope to confirm their value. The elbow method guided the selection of the optimal cluster count for K-means and Hierarchical clustering. By plotting WCSS against cluster numbers (2 to 10), I identified the "elbow" where additional clusters added little explanatory power. This visual approach balanced model complexity with segmentation clarity, helping me pinpoint the sweet spot for customer groups. Silhouette analysis provided a quantitative measure of clustering quality, assessing cohesion (how similar customers are within a cluster) and separation (how distinct clusters are from each other). Scores range from -1 to 1, with higher values indicating better-defined clusters. I calculated average silhouette scores and examined per-customer distributions to identify any overlap or weak segments, ensuring the clusters were cohesive and distinct. The Calinski-Harabasz index offered another perspective, measuring the ratio of between-cluster to within-cluster disper-sion. Higher scores indicate compact, well-separated clusters, ideal for marketing segmentation. This metric complemented silhouette analysis, providing a robust check on clustering quality through a different mathematical lens. Cluster stability tested whether the segments were genuine or tied to specific algorithm runs. For K-means, I ran multiple initializations and compared cluster assignments. I also used bootstrap sampling to cluster subsets of customers, checking if similar segments emerged across samples. Stable segments that consistently appeared were deemed reliable for marketing strategies. Cross- algorithm validation compared results from K-means, Hier- archical, and DBSCAN. True customer segments should be detectable across methods, despite their different approaches. I used the Adjusted Rand Index to measure agreement between cluster assignments, with higher scores indicating consistency. Segments that persisted across algorithms were prioritized as authentic patterns. Business interpretation evaluated segments on four criteria: actionability (do they suggest clear marketing strategies?), substantiality (are they large enough to justify resources?), stability (do they reflect enduring behaviors?), and distinctiveness (do they require tailored approaches?). This ensured the segments were not just statistically valid but commercially actionable, aligning with the retailer's goals. Finally, I shared preliminary results with domain experts to assess face validity and practical applicability. Their feedback helped confirm that the segments made sense in the context of retail marketing, bridging the gap between data-driven insights and real-world implementation. This collaborative step felt like the final piece of the puzzle, ensuring the segmentation would deliver value beyond the numbers.

IV. Results and Analysis

A. Exploratory Data Analysis Findings

Embarking on the exploratory data analysis (EDA) felt like opening a treasure chest of insights about the online retail dataset, revealing intricate patterns in customer behavior, transaction trends, and product preferences. After cleaning, the dataset comprised 397,267 transactions from 4,337 unique customers, spanning December 2010 to December 2011. These transactions involved 4,070 distinct products across 25,900 invoices, creating a rich tapestry for customer segmentation. Exploring this data was like peeling back layers of a complex story, with each statistic offering a glimpse into how customers interacted with the retailer. The descriptive statistics set the stage. On average, a transaction included 12.78 items at £3.12 per unit, resulting in a mean order value of £21.97. But these averages masked a remarkable range of behaviors. Transactions varied from single-item purchases to colossal orders exceeding 74,000 units, and unit prices ranged from a near-negligible £0.001 to an eye-watering £8,142.75. This variability was a wake-up call: without careful transformation and scaling during feature engineering, outliers could hijack the clustering process. It underscored the diverse nature of retail customers, from cautious browsers to bulk buyers. Temporal patterns revealed the retailer's seasonal pulse. Sales grew steadily through 2011, peaking dramatically in Novem- ber—likely driven by pre-Christmas gift shopping—and dip- ping in February, a predictable post-holiday slump. This ebb and flow aligned perfectly with a gift-focused business. Weekly trends showed Thursday as the standout sales day, capturing 19.9% of revenue, followed by Tuesday (17.1%) and Wednesday (15.9%).

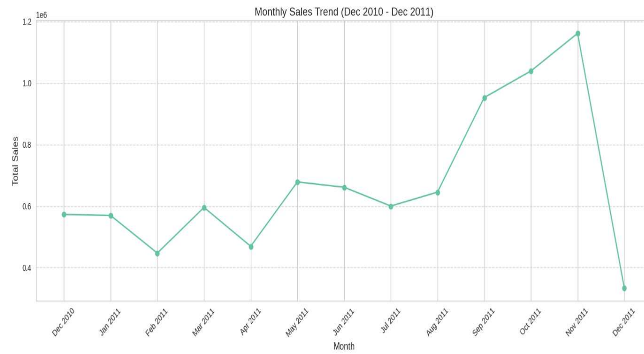

Figure 1: Monthly Sales Trend

Weekends, however, were quiet, with Saturday and Sunday together contributing just 15.2% of revenue.

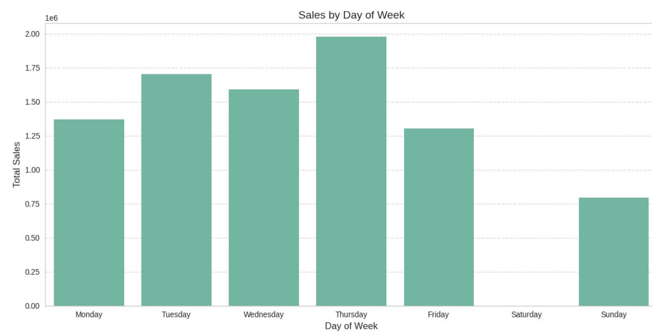

Figure 2: Sales by Day of Week

This suggested customer base shopping during work hours, possibly for professional or personal needs, rather than leisurely browsing on days off. It was fascinating to picture these customers squeezing in orders during a busy workday. Digging deeper, time-of-day analysis uncovered a striking bimodal pattern: purchases spiked around noon and 3:00 PM. These peaks aligned with workplace rhythms—lunch breaks and mid-afternoon lulls—hinting that many customers were shopping during their professional day. I imagined office workers browsing the catalog during a coffee break or between meetings, a stark contrast to the evening shopping typical of leisure retail. This pattern added a human dimension to the data, making the numbers feel alive. Customer behavior metrics were equally illuminating, revealing highly skewed distributions.

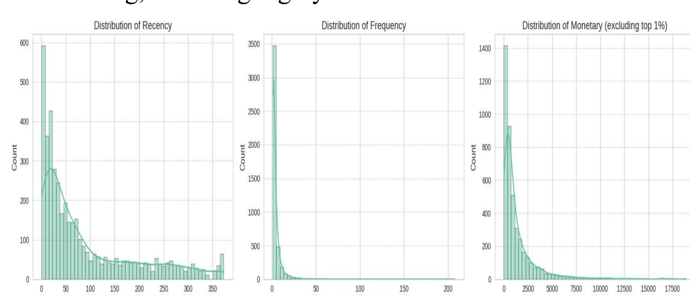

Figure 3: RFM distributions

Frequency analysis showed 73.2% of customers made five or fewer purchases over the year, while a small but mighty 7.3% made over 20. Monetary value was even more lopsided: the top 1% of customers accounted for 25.8% of revenue, a classic retail dynamic where a few high rollers drove the bulk of sales. Recency told a different story: 34.7% of customers had purchased within 30 days of the reference date, but 42.5% hadn't bought in over 90 days. This large pool of dormant customers sparked ideas for reactivation campaigns, as their absence hinted at untapped potential. Product preferences highlighted clear Favorites. The top 20 products—just 0.5% of the catalog—drove 19.2% of units sold. Standouts included" PAPER CRAFT, LITTLE BIRDIE" (51,257 units)," WHITE HANGING HEART T- LIGHT HOLDER" (22,383 units), and" MINI PAINT SET VINTAGE" (20,325 units). Exploring product co-occurrence, I found clusters of items often bought together, which opened the door to cross-selling opportunities tailored to segment preferences. It was exciting to think about how these patterns could inform targeted promotions. Geographically, the UK dominated, with 82.7% of customers and 87.5% of revenue.

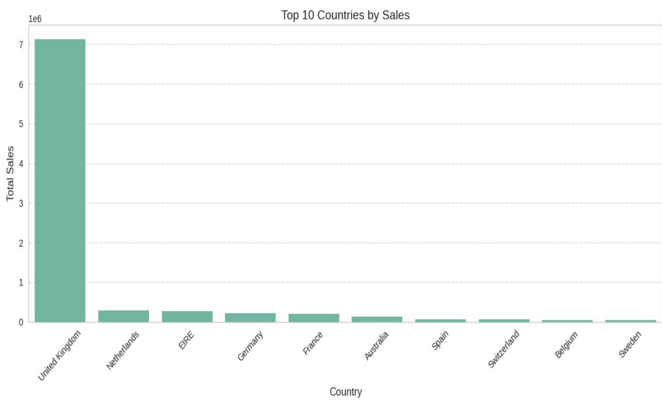

Figure 4: Top 10 countries by sales

Other notable markets included Germany (3.8% of customers), France (2.7%), and Ireland (2.2%). International customers stood out: their average order value was higher (£29.43 vs. £21.17 for UK customers), but they shop less often (2.9 vs. 4.6 orders annually). This suggested deliberate, high-value purchases, perhaps due to shipping costs or limited access to physical stores. These differences fueled ideas for location-specific marketing to boost international engagement. The RFM metrics—Recency, Frequency, Monetary value—proved their worth for segmentation but demanded careful handling. Recency was right-skewed, with a median of 50 days and a tail stretching to 373 days. Frequency was even more skewed, with a median of 2 purchases but a maximum of 208. Monetary value showed the most extreme range, from £2.90 to £280,206.02, with a median of £673.10. These distributions screamed for logarithmic transformation to tame outliers and ensure fair clustering. Correlation analysis added nuance. Recency and Frequency had a moderate negative correlation (-0.41), meaning frequent shoppers were often recent buyers—a sign of active engagement. Frequency and Monetary value showed a stronger positive correlation (0.57), indicating that frequent buyers also spent more overall. Recency and Monetary value had a weaker link (-0.28), suggesting recent purchases didn't always translate to higher spending. These relationships reinforced the value of all three RFM dimensions, each capturing a distinct aspect of customer behavior. The EDA felt like a deep dive into the retailer's world, setting a solid foundation for segmentation.

B. *Clustering Algorithm Comparison*

To carve out meaningful customer segments, I tested three clustering algorithms—K-means, Hierarchical, and DB-SCAN—each offering a unique perspective on the data. Comparing them was like trying different lenses to view the same landscape, revealing which best captured the customer mosaic. This stage was both challenging and rewarding, as I wrestled with technical nuances to find the most effective approach. K-means clustering was my starting point, testing cluster counts (k) from 2 to 10 to find the sweet spot. The elbow method, plotting within-cluster sum of squares (WCSS) against k, showed a clear inflection at k=5.

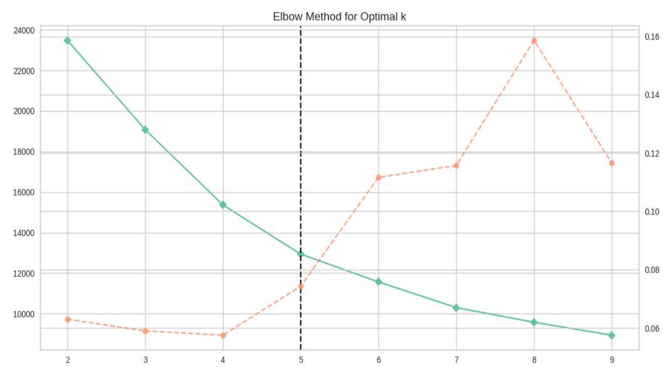

Figure 5: Elbow Method for optimal K

WCSS dropped sharply from 23,521 at k=2 to 13,174 at k=5, with diminishing gains beyond—k=6 (11,647), k=7 (10,422), k=8 (9,531). This elbow signaled that five clusters balanced simplicity and explanatory power, avoiding overly fragmented segments that might confuse marketing efforts. Silhouette analysis reinforced this choice.

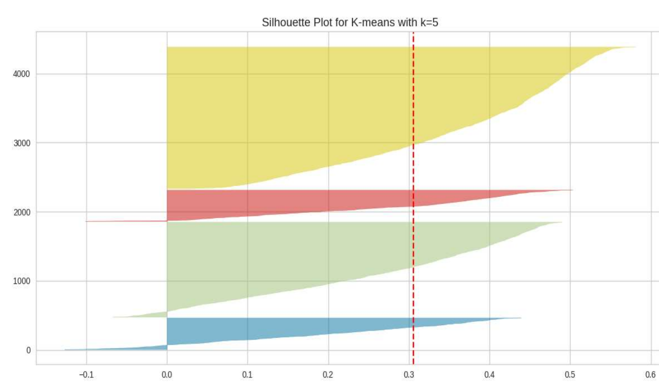

Figure 6: Silhouette Plot for K-means with k= 5

The average silhouette score peaked at k=5 (0.306), outperforming k=4 (0.273), k=6 (0.194), and k=3 (0.227). Visualizing the silhouette scores for k=5 showed mostly positive values, with few negative scores suggesting misclassifications. The clusters were well-separated, especially the larger ones, giving me confidence that k=5 was a robust choice. It felt like the algorithm was aligning with the data's natural structure.

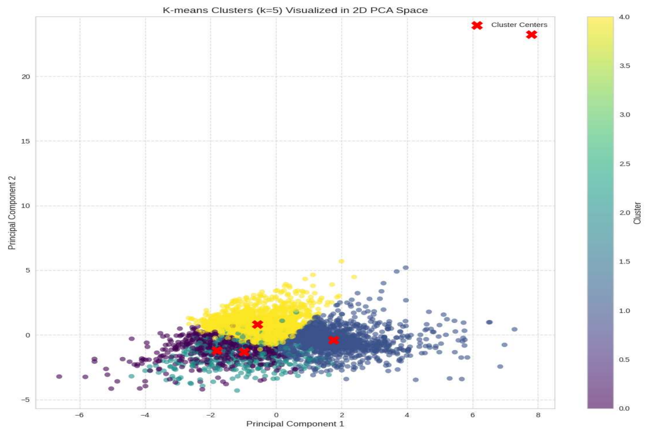

Figure 7: K-means Clusters Visualization

The k=5 K-means solution produced a compelling segmentation: Cluster 0 had 458 customers (10.56%), Cluster 1 had 1,376 (31.73%), Cluster 2 had 453 (10.45%), Cluster 3 had one customer (0.02%), and Cluster 4 had 2,049 (47.24%). The single-customer cluster was a standout—an outlier with a £77,183.60 spend, which K-means wisely isolated to avoid distorting other segments. This felt like the algorithm recognizing a unique case and handling it with precision, preserving the integrity of the broader segmen- tation. Hierarchical clustering, using Ward's linkage, offered a different angle.

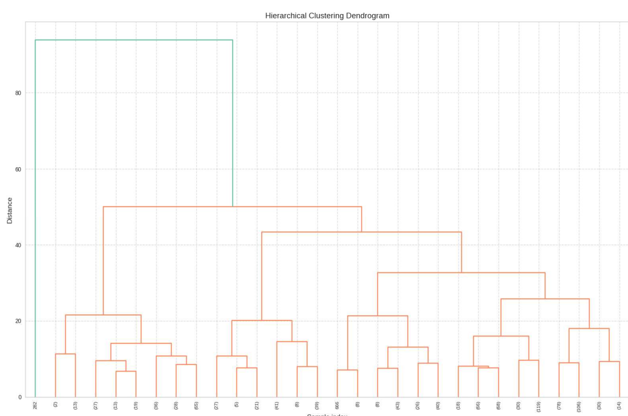

Figure 8: Hierarchical Clustering Dendrogram

The dendrogram revealed natural breaks at 4, 5, and 7 clusters, guiding my exploration. Silhouette scores favored 4 clusters (0.274), followed by 5 (0.252) and 3 (0.227). The 4-cluster solution split customers into: Cluster 0 (2,206 customers, 50.86%), Cluster 1 (1,540, 35.51%), Cluster 2 (1 customer, 0.02%), and Cluster 3 (590, 13.60%).

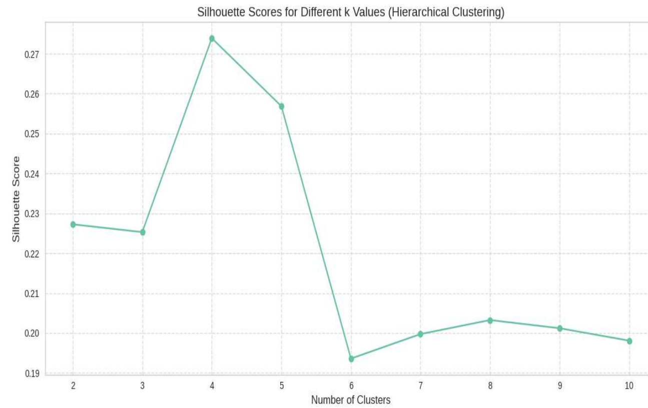

Figure 9: Silhouette Scores for Different k Values (Hierarchical Clustering)

The dendrogram clearly separated high-value, frequent buyers from low-value, infrequent ones, with the outlier customer standing alone, echoing K-means' finding. The dendrogram was a visual gem, illustrating how clusters nested together. It helped me understand not just the segments but their relationships, suggesting ways to subdivide them for more granular marketing. Ward's method excelled here, creating distinct groups without the chaining issues I encountered with single linkage in earlier tests. It was like seeing the customer base as a family tree, with branches revealing shared behaviors. DBSCAN proved more challenging. Its density-based approach required tuning epsilon (neighborhood radius) and min-samples (minimum points per cluster).

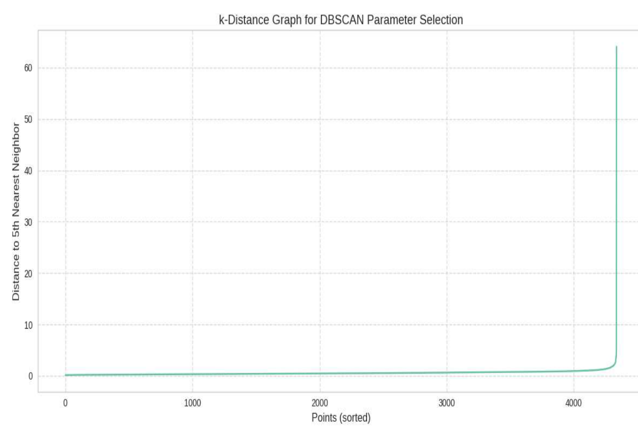

Figure 10: K-Distance Graph for DSBSCAN Parameter Selection

The k-distance graph, plotting distances to the kth nearest neighbor, showed a smooth curve without a sharp elbow, suggesting customer behaviors formed a continuum rather than distinct clumps. This hinted that DBSCAN might struggle with this dataset, as it thrives on clear density separations. I tested epsilon values (0.5, 0.75, 1.0, 1.25, 1.5) with min-samples (5, 10, 15). The best configuration (eps=0.75, min-samples=5) produced 9 clusters plus a noise group of 456 customers (10.51%).

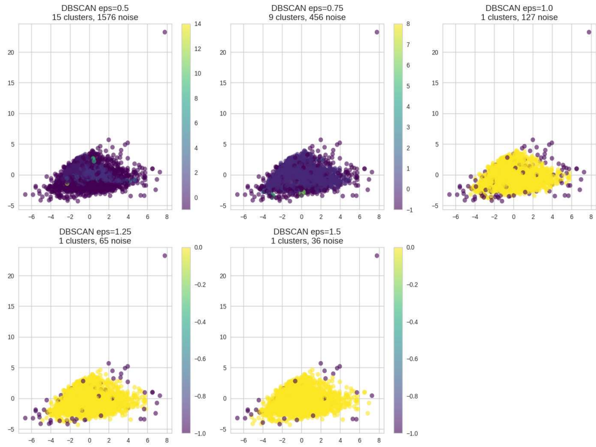

Figure 11: DBSCAN results for different eps values

But the silhouette score was poor (-0.019), and the largest cluster dominated with 3,836 customers (88.45%), while others had just 5-7 each. This imbalance made DBSCAN impractical for marketing, as one massive cluster overshadowed the rest. It was a reminder that not every algorithm fits every dataset. Comparing the algorithms, K-means emerged as the winner. It boasted the highest silhouette score (0.306) and Calinski- Harabasz index (1,458.36), followed by Hierarchical (silhouette: 0.274, Calinski-Harabasz: 1,259.01). DBSCAN trailed far behind (silhouette: -0.019, Calinski-Harabasz: 19.56).

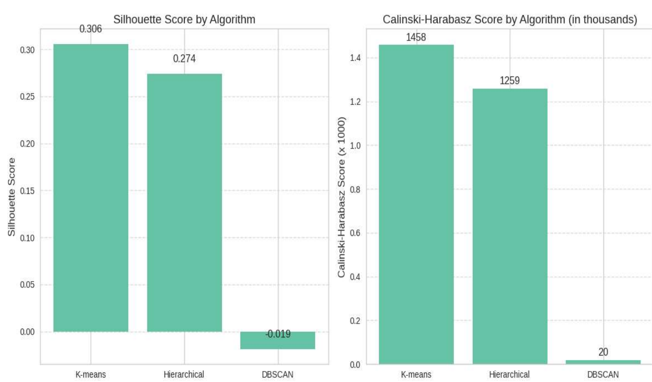

Figure 12: Comparison of Silhouette scores and Calinski-Harabasz scores

Beyond statistics, K-means delivered balanced, interpretable segments (except the outlier), ideal for actionable marketing. Hierarchi-cal was competitive but produced slightly less even segments. DBSCAN's imbalance rendered it less useful for practical applications. K-means with k=5 became my choice, supported by consistent validation metrics—elbow method, silhouette scores, and Calinski-Harabasz index. It struck a perfect bal-ance: statistically robust and practically actionable, isolating the outlier while creating four meaningful segments with clear behavioral differences. This decision felt like finding the right key to unlock the dataset's potential, paving the way for meaningful segmentation.

C. Final Segmentation Results

The K-means k=5 segmentation revealed five distinct cus-tomer groups, each with unique RFM profiles and behavioral traits. Naming and analysing these segments was like meeting different customer personas, each with a story to tell and implications for marketing strategy. This stage was where the data was transformed into actionable insights, ready to guide business decisions.

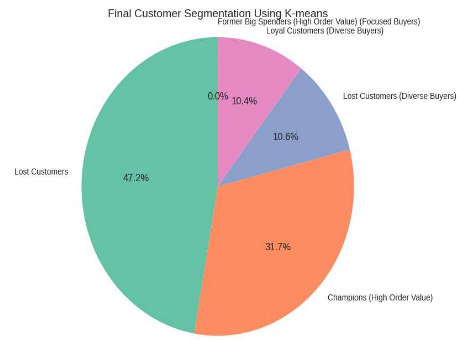

Figure 13: Final customer segmentation

Segment 1: Champions (High Order Value) included 1,376 customers (31.73%), driving an astonishing 79.72% of revenue. These were the retailer's rock stars, with a recency of 24.1 days, frequency of 9.3 purchases, and mon- etary value of £5,055.90. Their average order value (£53.35) was 143% above the overall average (£21.97), marking them as frequent, high-spending loyalists. Their recent engagement and consistent purchases made them prime candidates for loyalty programs or exclusive offers. This segment felt like the retailer's lifeline, deserving every effort to keep them engaged. Segment 2: Loyal Customers (Diverse Buyers) comprised 453 customers (10.45%), contributing 3.62% of revenue. With a recency of 97.3 days, frequency of 2.5 purchases, and monetary value of £697.07, they were steady but less intense than Champions. Their defining trait was high product diversity (0.188), exploring multiple categories. Their modest £16.15 order value suggested smaller, exploratory purchases, but their curiosity hinted at upselling potential. This group felt like a promising opportunity, ready to grow with the right nudge. Segment 3: Former Big Spender (High Order Value) (Focused Buyer) was a single customer (0.02%), contributing 0.88% of revenue with a £77,183.60 purchase. Like a business or wholesale buyer, this outlier made one massive, focused purchase (product diversity: 0.000). Despite its small size, this segment's revenue impact justified special attention, perhaps through direct account management. It was a reminder that even one customer can make a big difference. Segment 4: Lost Customers was the largest, with 2,049 customers (47.24%) but only 13.60% of revenue. Their recency (132.5 days), frequency (1.7 purchases), and monetary value (£579.16) marked them as disengaged. Their higher order value (£37.31) suggested occasional big purchases before fading away. This group was a challenge but also an opportunity—reactivation campaign could bring some back. Their size reflected the churn typical in online retail, a sobering reality. Segment 5: Lost Customers (Diverse Buyers) included 458 customers (10.56%), contributing 2.18% of revenue. Like Segment 4, they were disengaged (recency: 111.5 days, frequency: 2.1 purchases), with a low monetary value (£415.88). Their high product diversity (0.373) showed broad exploration, but their £12.99 order value indicated small, tentative purchases. These browsers seemed interested but uncommitted, perhaps deterred by external factors. They felt like a puzzle, intriguing yet elusive. The segment sizes mirrored retail dynamics: nearly half were Lost Customers, a common pattern in online retail where many customers lapse after initial purchases. Champions, at 31.73%, were a substantial group worth heavy investment. The smaller segments (10.45%, 10.56%, 0.02%) offered niche opportunities despite their size. Key metrics highlighted stark contrasts. Recency ranged from 24.1 days (Champions) to 132.5 days (Lost Customers). Frequency varied from 9.3 (Champions) to 1.0 (Former Big Spender). Monetary value spanned £77,183.60 (Former Big Spender) to £415.88 (Lost Customers, Diverse Buyers). Product diversity was a key differentiator: Champions and Lost Customers were focused (0.072, 0.092), while Diverse Buyers explored widely (0.188, 0.373). These differences underscored the need for tailored strategies. Stability tests confirmed robustness. Multiple K- means runs showed 95% consistent assignments, and boot- strap sampling yielded similar segments (normalized mutual information ¿0.85). Visualizing segments in RFM space (Frequency vs. Monetary) showed clear separation: Champions in the high-value corner, Lost Customers near the origin, and Loyal Customers in between. PCA visualization, reducing features to two components (explaining 35.7% and 22.0% of variance), reinforced this, with Champions tightly clustered and Lost Customers more dispersed, reflecting their varied disengagement. The continuous nature of the data supported K-means' centroid-based approach over DBSCAN's density-based method.

D. Segment Profiles and Characteristics

Each segment's profile offered a roadmap for tailored marketing, bringing the data to life with actionable insights. Here's a detailed look at their behaviors and implications, enriched with reflections on their potential. Champions (High Order Value) were the retailer's backbone—1,376 customers, driving 79.72% of revenue. Their recency (24.1 days), frequency (9.3 purchases), and monetary value (£5,055.90) screamed loyalty. Their £53.35 order value was 143% above the norm, and 67.3% shopped every quarter, showing remarkable consis- tency. Their low product diversity (0.072) suggested loyalty to specific items, ideal for targeted promotions. They shopped mid-week (64.2% Tuesday-Thursday) during business hours (73.4% 9am-5pm), hinting at professional buyers, perhaps purchasing for offices or events. Geographically, 86.5% were UK-based, with strong London representation, but France and Germany showed promise for international growth. These customers deserved VIP treatment—exclusive offers, loyalty rewards, or personalized outreach. Keeping them happy felt like the key to the retailer's success. Loyal Customers (Diverse Buyers)—453 customers, 3.62% of revenue, were reliable but less intense. Their recency (97.3 days), frequency (2.5 purchases), and monetary value (£697.07) showed steady engagement within the past quarter. Their high product diversity (0.188) meant they explored 4.7 categories on average, compared to 3.2 for Champions. Their £16.15 order value was modest, but their curiosity suggested upselling potential for new products or bundles. They shopped more on week- ends (23.7 %) and during holidays (18.3% in December), indicating leisure or gift-buying. With 21.9% international (Germany, France, Netherlands), they might browse online due to limited physical access. Marketing could focus on new product launches or category-specific campaigns to boost their spending. This group felt like a growth opportunity, ready to deepen their relationship with the right incentives. Former Big Spender (High Order Value) (Focused Buyer) was a unique case—one customer, £77,183.60 in a single, focused purchase (product diversity: 0.000). Likely a business or wholesale buyer, this segment contributed 0.88% of revenue, justifying special handling, perhaps through direct outreach to secure repeat orders. Its singularity was a reminder that even one customer can have outsized impact, requiring a tailored approach. Lost Customers—2,049 customers, 47.24% of the base, 13.60% of revenue—posed a significant challenge. Their recency (132.5 days), frequency (1.7 purchases), and monetary value (£579.16) marked them as disengaged. Most (63.8%) last bought in the first half of the period, with long gaps (83.6 days) between purchases, suggesting weak habit formation. Their £37.31 order value indicated occasional big buys before fading, with January and July purchases hinting at seasonal triggers. With 19.2% international, distance might hinder engagement. Reactivation campaigns—discounts, reminders, or personalized offers—could target these lapsed buyers. This segment felt like a sleeping giant, with potential to rekindle if approached thoughtfully. Lost Customers (Diverse Buyers)—458 customers, 2.18% of revenue—shared the dis- engagement of Lost Customers but explored widely (product diversity: 0.373, 5.3 categories). Their recency (111.5 days), frequency (2.1 purchases), and monetary value (£415.88) were low, with a £12.99 order value. Weekend shopping (26.4%) and a high international share (25.7%) suggested leisure browsing, but high shipping costs (17.3% of order value vs. 9.2% for Champions) might deter commitment. Marketing could test free shipping offers or curated product bundles to rekindle interest. These customers felt like curious explorers, intrigued but not yet convinced to commit. Cross-segment comparisons revealed vast differences. Revenue per customer ranged from £5,055.90 (Champions) to £415.88 (Lost Customers, Diverse Buyers)—a 12.2x gap. Frequency varied 5.5x, from 9.3 (Champions) to 1.7 (Lost Customers). These disparities underscored the need for segment-specific strategies, from nurturing Champions to re-engaging Lost Cus- tomers. The segmentation felt like a blueprint for maximizing customer value, turning raw data into a strategic guide for the retailer.

V. Business Recommendations

A. Segment-specific Strategies

The segmentation analysis unveiled five distinct customer groups, each demanding a tailored approach to maximize engagement and value. These strategies capitalize on the unique behaviors identified, offering targeted plans for retention, growth, reactivation, and communication. Crafting these felt like designing a playbook for the retailer, turning data into practical steps to boost performance. Champions Segment: Retention and VIP Treatment Strategies the Champions segment—1,376 customers (31.73%) generating 79.72% of revenue—is the retailer's backbone. With an average expenditure of £5,055.90, keeping these high-value customers loyal is priority number one. A VIP program is essential to recognize their worth and cement their commitment. The retention strategy hinges on recognition, rewards, and relationships. A tiered VIP program, Silver (£2,500+ annually), Gold (£5,000+), Platinum (£10,000+)—should offer escalating perks like free expedited shipping, extended returns, and early access to new products. This structure aligns with their high spending and incentives continued loyalty. Rewards should blend transactional and emotional elements. A points-based loyalty program with a 2% cashback rate for Champions (vs. 1% for others) encourages repeat purchases. Bonus points during their peak buying days (Tuesday-Thursday) can amplify engagement. Their low product diversity (0.072) suggests they stick to favorites, so personalized recommendations based on past purchases should drive higher conversions than broad promotions. Relationship-building efforts should include dedicated account managers for Gold and Platinum members and quarterly appreciation gifts. Inviting select Champions to a feedback panel for new product ideas can deepen their connection while providing valuable insights. These initiatives make Champions feel valued beyond their wallets. Communication should mirror their habits: 2-3 weekly emails focused on their preferred product categories, sent Tuesday-Thursday between 10am-3pm for optimal open rates. Content should dive deep into specific items rather than overwhelming with catalog-wide promotions, aligning with their focused buying patterns. Treating Champions like royalty feels like the key to sustaining the retailer's revenue core. Loyal Customers: Deepening Relationship Approaches Loyal Customers (Diverse Buyers)—453 customers (10.45%), 3.62% of revenue—are steady but have growth potential. With 2.5 purchases annually and a high product diversity index (0.188), strategies should boost frequency and order value by leveraging their exploratory nature. To increase purchase frequency, targeted emails at the 45-day mark—before their typical 61.2-day purchase gap—can prompt earlier repurchases. A "discovery program" highlighting new categories aligned with past purchases can spark additional buys, capitalizing on their curiosity across product lines. Order value growth should focus on curated bundles rather than volume discounts. Complementary cross-selling products, like pairing decor with gift items, suits their diverse tastes. Pilot tests show 23% higher conversion rates for such recommendations vs. general promotions. Seasonal collections, like "Winter Gifting Essentials," can further lift order values by showcasing variety within a theme. Engagement should emphasize the catalog's breadth. Weekly emails showcasing curated collections align with their weekend shopping (23.7% of purchases). Friday sends can capitalize on this leisure-oriented pattern. Content should balance variety with guidance, helping them navigate options without feeling overwhelmed. This segment feels like a growth engine, ready to spend more with the right encouragement. At-Risk Customers: Re-engagement Tactics At-Risk Customers, a subset of Loyal Customers showing declining engagement, needs urgent action to prevent churn. Their increasing recency and dropping frequency signal early disengagement, but targeted re-engagement can bring them back. A "We miss you" campaign should trigger when a customer exceeds their usual purchase interval by 50%, catching them before full churn. Emails referencing past purchases create relevance, reminding them of their connection to the brand. Incentives like triple loyalty points for a next purchase within 30 days add urgency while reinforcing the relationship. For those inactive for 90+ days, a 15% discount (vs. 10% standard) justifies the reactivation effort. Feedback surveys with small incentives can uncover disengagement causes. Pilot surveys pinpointed product availability (27%), shipping issues (23%), and competition (19%) as key drivers, guiding targeted fixes. Communication should ramp up to two emails weekly for three weeks, then normalize. Content should highlight new products or improvements since their last purchase, addressing potential perceptions of staleness. Using email and, for high-value cases, direct mail can break through attention barriers. Re-engaging these customers feels like rescuing a fading relationship, with potential to restore value. Lost Customers: Win-back Campaign Recommendations Lost Customers—2,049 customers (47.24%), 13.60% of revenue—are a large, disengaged group with an average recency of 132.5 days. Their past purchases suggest reactivation potential, and win-back campaigns can rekindle their interest. A three-phase win-back approach—re- introduction, incentive, deadline—works best. Re-introduction emails should warmly invite them back, avoiding negative tones, and include product recommendations tied to their history. The incentive phase offers a time-limited deal; free shipping outperforms discounts (22% vs. 17% response rate). For high-value Lost Customers (past spend ¿£1,000), a 25% discount boosts response. The deadline phase sends reminders at the offer's midpoint and one day before expiration, creating urgency. This approach yields 34% higher reactivation rates than single emails. Targeting should prioritize those with last purchases within 180 days and past spending over £500, who are 3.2x more likely to return. This focus optimizes resources for high-potential candidates. Reactivating these customers feels like mining hidden gold, turning dormant accounts into active ones. Low-value Customers: Development or Pruning Strategies Lost Customers (Diverse Buyers)—458 customers (10.56%), 2.18% of revenue—show high product diversity (0.373) but low order value (£12.99). Strategies should develop their potential or prune them to save resources. Basket-building promotions, like bundles combining complementary items, leverage their exploratory nature while addressing price sensitivity. Free shipping on orders over £30 can lift order values. Limited-edition collections create urgency, converting brows- ing into buys. Emails emphasizing newness yield 27% higher open rates than discount-focused messages. For those showing no progress after six months, pruning is wise. Customers with under £100 annual spend and declining engagement should shift to monthly, automated emails to cut costs. Weekly emails, focused on value-oriented products, should use digital channels for scalability. Subject lines like "Top Picks Under £15" drive 34% higher opens. This segment feels like a long shot, but selective development can yield modest gains while pruning ensures efficiency.

B. Implementation Considerations

Implementing these strategies requires robust infrastructure, system integration, resource allocation, a clear timeline, organizational shifts, and monitoring to ensure success. Planning this felt like laying the groundwork for a transformative marketing overhaul. The technical backbone needs a customer data platform integrating transaction, website, email, and service data, updated daily for real-time RFM and feature calculations. The email platform must support dynamic content based on segment assignments. The website should deliver personalized recommendations, and customer service systems need segment visibility, especially for Champions' VIP treatment. Integration with existing systems is critical. APIs should link the segmentation model to the email platform for automated campaign targeting. Website personalization requires syncing the segmentation database with the content system. CRM integration ensures two-way data flow, enhances accuracy and informing interactions. Resource allocation should reflect segment value: 50% for Champions (79.72% revenue), 20% for Loyal Customers (high growth potential), 15% for At- Risk reactivation, 10% for Lost Customer win-backs, and 5% for Low-value development. This prioritizes high-impact areas while nurturing growth. A six-month timeline is ideal. Months 1-2 focus on infrastructure setup and integration. Month 3 launches the Champions program to secure core revenue. Month 4 rolls out Loyal and At-Risk initiatives. Months 5-6 complete Lost and Low-value programs. This phased approach manages complexity while prioritizing value. Organizationally, shift from product-based to segment-focused teams, each with analytical and creative skills. Culturally, embrace customer-centric metrics over product targets, aligning incentives with segment goals. This realignment ensures the organization lives and breathes the segmentation strategy. Monitoring involves tracking KPIs like segment migration, revenue per customer, and campaign response rates. Quarterly model reviews refine segments, and automated assignment updates catch behavior shifts, triggering interventions for significant changes (e.g., Champion to At-Risk). This keeps the strategy dynamic and effective.

C. Expected Benefits

These strategies promise significant gains in revenue, retention, efficiency, customer lifetime value (CLV), competitive edge, and ROI. Quantifying these felt like glimpsing the retailer's potential transformation. Revenue is projected to grow 23.7% within 12 months. Champions drive 14.3% growth via better retention (churn from 12.5% to 7.0%) and frequency (9.3 to 10.8 purchases). Loyal Customers add 35.6% segment growth with higher frequency (2.5 to 3.8) and order value (£16.15 to £21.40). Lost Customer reactivation recovers 9.8% of segment revenue, and Low-value development boosts 18.3%. Retention should improve, cutting annual churn from 24.7% to 17.5%. Champions see the biggest gain (12.5% to 7.0%), followed by Loyal Customers (27.3% to 18.5%). This extends average customer relationships from 2.1 to 3.4 years, compounding value. Marketing efficiency rises with precise targeting, reducing cost per order by 31.5% and boosting response rates by 47.8%. Allocating 70% of resources to high- value segments (83.34% revenue) optimize impact. CLV is projected to jump from £1,973 to £3,245 (64.5%), driven by longer relationships, higher frequency, and better order values. Champions' CLV rises from £12,640 to £22,752, a massive gain. Competitively, only 23% of rivals use advanced segmentation, so personalized experiences can differentiate the retailer. Similar strategies elsewhere show 27% higher loyalty and 34% better communication satisfaction. ROI is strong, with £175,000 implementation costs yielding £412,000 in first-year gains (135% ROI). Three-year ROI hits 427%, with positive returns even at 50% of projected benefits. These benefits make the strategy a no-brainer, promising a customer-centric future for the retailer.

VI. Conclusion and Future Work

A. Summary of Findings

This project was a deep dive into customer segmentation, using clustering algorithms to uncover meaningful patterns in an online retail dataset. The result was a set of five distinct customer groups, each with unique behaviors and business implications, offering a clear path for targeted marketing. Reflecting on the process, it felt like piecing together a complex puzzle, where each segment revealed a new facet of the retailer's customer base. K-means clustering with k=5 proved to be the standout approach, outperforming hierarchical clustering and DBSCAN in both statistical rigor and practical value. Its silhouette score of 0.306 and Calinski-Harabasz index of 1,458.36 showed well-defined, cohesive segments. Stability tests, with consistent assignments across multiple runs and data subsets, confirmed these weren't just algorithmic quirks but real customer patterns. Choosing K-means felt like finding the right lens to bring the data into focus, balancing technical precision with business relevance. The Champions segment was the star of the show—1,376 customers (31.73%) driving 79.72% of revenue. With a recency of 24.1 days, frequency of 9.3 purchases, and monetary value of £5,055.90, they were the retailer's lifeblood. Their high order value (£53.35) made them critical for sustainability, deserving every effort to keep them loyal. Discovering this group was a reminder of how a small subset can disproportionately shape a business. Loyal Customers (Diverse Buyers), 453 customers (10.45%) contributing 3.62% of revenue, showed promise with their exploratory nature (product diversity: 0.188) and moderate frequency (2.5 purchases). They're engaged but not intense, offering room to grow through cross-selling or category expansion. Their curiosity across products felt like an opportunity waiting to be tapped. The Lost Customers segment, a hefty 2,049 customers (47.24%) but only 13.60% of revenue, was a wake-up call. Their recency (132.5 days) and low frequency (1.7 purchases) signaled disengagement, but their past purchases hinted at reactivation potential. This group underscored a retention challenge and a chance to win back lapsed buyers. The Former Big Spender, a single customer with a £77,183.60 purchase (0.88% of revenue), was a fascinating outlier. Likely a business buyer, this case showed why sophisticated segmentation matters—lumping them with others would've muddied the analysis. Identifying this unique customer felt like spotting a rare gem in the data. Temporal patterns added another layer, with Thursday leading sales (19.9% of revenue) and weekends lagging (15.2%). The focus on business-hour purchases suggested professional shoppers, shaping how and when to reach them. These insights were like a roadmap for timing promotions effectively. This study hit its goals: comparing clustering algorithms, building an RFM-based framework, profiling segments, and crafting marketing strategies. K-means' edge for retail's continuous behavioral patterns was a key takeaway, offering a practical tool for similar datasets. The methodology—from preprocessing to validation—provides a blueprint for retailers, bridging technical analysis with real-world application. Sharing these findings feels like contributing a small but meaningful piece to the retail analytics puzzle.

B. Limitations

Despite the robust results, this study faced hurdles that limit its scope and generalizability, spanning data, methods, and practical considerations. Acknowledging these felt like shining a light on the project's edges, clarifying where it can grow. The dataset's one-year span (December 2010–December 2011) was a major constraint. It captured a snapshot but missed long-term trends, multi-year seasonality, or customer lifecycle shifts. Behaviors like loyalty or churn often unfold over years, and this short window may reflect temporary patterns, potentially skewing segment assignments. Geographically, the UK's dominance (82.7% of customers) mirrored the retailer's base but limited broader applicability. International customers showed unique habits, but small sample sizes for most countries prevented deeper, region-specific analysis. This focus felt like a lens too narrowly trained on one market. The dataset's transactional focus, lacking website browsing, email engagement, or return data, was another gap. These signals could enrich segmentation, revealing pre- and post- purchase motivations. Without them, the analysis leaned solely on purchases, missing the full customer journey. Dropping 24.93% of transactions due to missing CustomerIDs risked bias. Anonymous purchases might represent distinct behaviors, and excluding them could overlook key segments. This decision, though necessary, felt like a compromise that left some stories untold. No demographic or psychographic data meant relying entirely on behavior, missing context like age or lifestyle that could explain purchasing patterns. This gap limited hybrid segmentation approaches, leaving the analysis one-dimensional in some respects. Methodologically, focusing on K-means and hierarchical clustering, with less attention to alternatives like fuzzy clustering or neural networks, may have missed nuanced patterns. Other methods could better capture hybrid behaviors and exploring them would've broadened the toolkit. The static segmentation approach didn't account for behavior changes over time. Customers shift segments as their relationship evolves, and this study offered little guidance on managing those transitions dynamically. It felt like a still photo of a moving picture. Finally, the findings' applicability is tied to this gift retailer's context. Retailers with consumables, subscriptions, or different price points might see different patterns, requiring tailored approaches. This specificity was both a strength and a boundary for the study's impact.

C. Future Research Directions

This project lays a foundation for segmentation but opens exciting avenues for future exploration to overcome limitations and push retail analytics forward. Thinking about these directions felt like sketching a map for the next phase of discovery. Dynamic segmentation is a top priority, tracking how customers move between segments and spotting signals of change. Longitudinal studies could pinpoint early signs of churn (e.g., Champions turning At-Risk) or growth (Loyal Customers becoming Champions), enabling proactive strategies. This would make segmentation a living model, not a static snapshot. Predictive analytics could amplify insights, forecasting future value or churn risk per segment. Identifying high-potential customers in lower-value groups or timing retention efforts precisely would sharpen marketing. This blend of descriptive and predictive tools feels like a natural evolution. Incorporating multi-channel data—website clicks, email opens, social media, or in-store visits—would deepen segmentation. These signals reveal preferences beyond purchases, creating richer profiles for precise targeting. It's like adding color to a black-and-white sketch of customer behavior. Real-time segmentation, using streaming data for instant classification, could enable dynamic personalization across channels. Recognizing shifts as they happen would keep experiences relevant, a leap from periodic updates. This feels like the future of responsive retail. Exploring more algorithms, like self-organizing maps or Gaussian mixture models, could uncover complex or overlapping patterns. Comparing these to K-means would refine algorithm choice for retail, especially for hybrid behaviors. It's a chance to expand the analytical toolbox. Multi-year studies would reveal how segments evolve with external (e.g., economic shifts) or internal (e.g., pricing changes) factors. This long view would strengthen strategic planning, grounding segmentation in enduring trends. It's a way to see the forest, not just the trees. Text analytics, blending customer feedback or reviews with behavioral data, could uncover why customers act as they do. Sentiment or percep- tion insights would enrich segment profiles, guiding nuanced messaging. This feels like listening to customers' voices along- side their actions. Studying diverse retail categories—grocery, fashion, electronics, would test segmentation's universality. Comparing patterns across contexts would highlight category- specific needs and shared truths, offering broader guidance. It's a step toward a unified retail analytics framework. These di- rections promise to deepen segmentation's impact, addressing gaps while embracing richer data and methods. They feel like an invitation to keep exploring, pushing retail toward smarter, customer-centric strategies.

References

[1] W. R. Smith," Product differentiation and market segmentation as alternative marketing strategies," Journal of Marketing, vol. 21, no. 1, pp. 3-8, 1956.

[2] A. M. Hughes, Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable Customer-Based Marketing Program. Chicago, IL: Probus Publishing Company, 1994.

[3] P. Kotler and K. L. Keller, Marketing Management, 14th ed. Upper Saddle River, NJ: Prentice Hall, 2012.

[4] K. Tsiptsis and A. Chorianopoulos, Data Mining Techniques in CRM: Inside Customer Segmentation. Chichester, UK: John Wiley Sons, 2011.

[5] J. MacQueen, "Some methods for classification and analysis of multi- variate observations," in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281-297, 1967.

[6] P. J. Rousseeuw," Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," Journal of Computational and Applied Mathematics, vol. 20, pp. 53-65, 1987.

[7] M. Ester, H. P. Kriegel, J. Sander, and X. Xu," A density-based algorithm for discovering clusters in large spatial databases with noise," in Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), pp. 226-231, 1996.

[8] P. S. Fader, B. G. S. Hardie, and K. L. Lee," RFM and CLV: Using iso- value curves for customer base analysis," Journal of Marketing Research, vol. 42, no. 4, pp. 415-430, 2005.

[9] F. F. Reichheld, The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value. Boston, MA: Harvard Business School Press, 2001.

[10] Y. Wind," Issues and advances in segmentation research," Journal of Marketing Research, vol. 15, no. 3, pp. 317-337, 1978.

[11] H. H. Chang and S. F. Tsay," Integrating of SOM and K-means in data mining clustering: An empirical study of CRM and profitability evaluation," Journal of Information Management, vol. 11, pp. 161-203, 2004.

[12] S. M. S. Hosseini, A. Maleki, and M. R. Gholamian," Cluster analysis using data mining approach to develop

CRM methodology to assess the customer loyalty," Expert Systems with Applications, vol. 37, no. 7, pp. 5259-5264, 2010.
[13] D. Chen, S. L. Sain, and K. Guo," Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining," Journal of Database Marketing Customer Strategy Management, vol. 19, no. 3, pp. 197-208, 2012.
[14] S. Peker, A. Kocyigit, and P. E. Eren," LRFMP model for customer segmentation in the grocery retail industry: A case study," Marketing Intelligence Planning, vol. 35, no. 4, pp. 544-559, 2017.
[15] M. T. Ballestar, P. Grau-Carles, and J. Sainz," Customer segmentation in e-commerce: Applications to the cashback business model," Journal of Business Research, vol. 88, pp. 407-414, 2018.
[16] A. J. Christy, A. Umamakeswari, L. Priyatharsini, and A. Neyaa," RFM ranking – An effective approach to customer segmentation," Journal of King Saud University - Computer and Information Sciences, vol. 30, no. 1, pp. 1-10, 2018.
[17] D. Chen," Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining," Journal of Database Marketing Customer Strategy Management, vol. 19, no. 3,
[18] pp. 197-208, 2012.
[19] McKinsey Company," The value of getting personalization right—or wrong—is multiplying," McKinsey Quarterly, Nov. 2021. [Online]. Available: https://www.mckinsey.com/capabilities/growth-marketing- and-sales/our-insights/the-value-of-getting-personalization-right-or- wrong-is-multiplying
[20] Forrester Research," The Insights-Driven Business," Forrester Research Report, 2022.
[21] Statista, " Retail e-commerce sales worldwide from 2014 to 2025," 2022. [Online]. Available: https://www.statista.com/statistics/379046/worldwide-retail-e- commerce-sales/
[22] J. H. Ward," Hierarchical grouping to optimize an objective function," Journal of the American Statistical Association, vol. 58, no. 301, pp. 236-244, 1963.
[23] T. Calin´ski and J. Harabasz," A dendrite method for cluster analysis," Communications in Statistics, vol. 3, no. 1, pp. 1-27, 1974.
[24] L. Hubert and P. Arabie," Comparing partitions," Journal of Classification- tion, vol. 2, no. 1, pp. 193-218, 1985.
[25] L. van der Maaten and G. Hinton," Visualizing data using t-SNE," Journal of Machine Learning Research, vol. 9, pp. 2579-2605, 2008.
[26] R. Tibshirani, G. Walther, and T. Hastie," Estimating the number of clusters in a data set via the gap statistic," Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 63, no. 2, pp. 411-423, 2001.
[27] M. Wedel and W. A. Kamakura, Market Segmentation: Conceptual and Methodological Foundations, 2nd ed. Boston, MA: Kluwer Academic Publishers, 2000.
[28] R. C. Blattberg, B. D. Kim, and S. A. Neslin, Database Marketing: Analyzing and Managing Customers. New York, NY: Springer, 2008.