International Journal of Science, Strategic Management and Technology
Volume **02** Issue **02** February-2026 | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Improving Retail Strategies Through Customer Segmentation to Find Target Audience

A.V V Satyanarayana
dept of Artificial Intelligence & Data Science
Koneru Lakshmaiah Education Foundation Vaddeswaram,India
M. Santhosh kumar
dept of Artificial Intelligence & Data Science
Koneru Lakshmaiah Education Foundation Vaddeswaram,India
R. Gowtham
dept of Artificial Intelligence & Data Science
Koneru Lakshmaiah Education Foundation Vaddeswaram,India
V. Tanishq
dept of Artificial Intelligence & Data Science
Koneru Lakshmaiah Education Foundation Vaddeswaram,India
Dr. Venkateshwarlu Sunkari
dept of Artificial Intelligence &Data Science
Koneru Lakshmaiah Education Foundation Vaddeswaram,India

https://doi.org/10.55041/ijsmt.v2i2.043

Cite this Article: Tanishq, A. V. S. ,. M. S. K. ,. R. G. ,. V. (2026). Improving Retail Strategies Through Customer Segmentation to Find Target Audience. International Journal of Science, Strategic Management and Technology, *Volume 10*(01). https://doi.org/10.55041/ijsmt.v2i2.043

License: This article is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting use, distribution, and reproduction in any medium, provided the original author(s) and source are properly credited.

***Abstract*—** In today's rapidly changing businesses industry, companies need to understand how technical market segmentation can give them a competitive advantage to target the right customer. To attain this competitive advantage, companies can employ market segmentation as a strategic approach, dividing the broader market into smaller, more homogeneous consumer groups who share similar characteristics. These characteristics encompass demographics like age, gender, income, education, and marital status, geographic location like town, city, state, or country, and behavior. The practice of market segmentation is vital for businesses as it enables them to focus their marketing efforts and product development more precisely. This paper explores the significance of technical market segmentation in aligning product development with market demands. Through a comprehensive review and analysis of a case study, this paper aims to provide valuable insights into effective segmentation strategies and their practical implementation for businesses.

I. Introduction

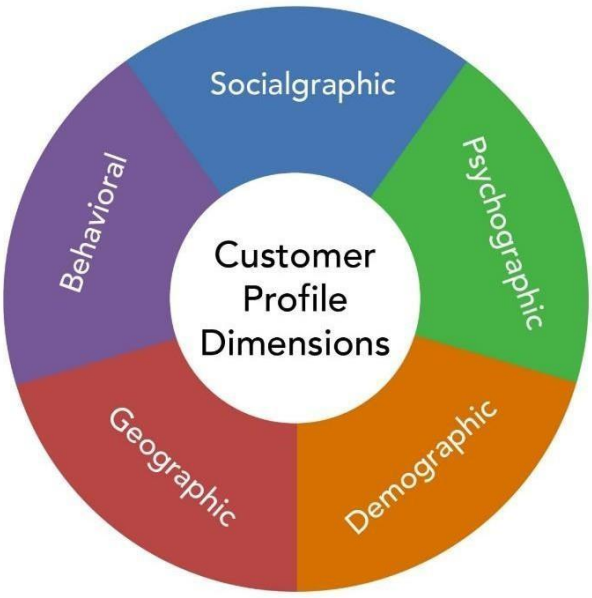
The competitive retail landscape today demands innovative ways for businesses to connect with their audiences and achieve success [1]. The era of generalized marketing is over, as modern consumers expect personalized, tailored experiences. This shift has transformed how companies approach market segmentation—the process of dividing a broad market into more focused, homogeneous groups based on shared traits [2]. Effective segmentation helps companies precisely target marketing and development efforts, leading to higher sales, customer satisfaction, and profitability.Traditionally, businesses have segmented customers using market research, surveys, and demographic profiling. While informative, these approaches often miss key aspects of consumer

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

behavior and preferences, especially in today's data-rich world. Their limitations include susceptibility to biases, difficulty adapting to market changes, and underuse of data resources [3]. Machine learning algorithms like K-means clustering address these gaps. As a powerful, unsupervised learning technique, K-means clustering analyzes complex datasets to reveal hidden patterns, categorizing customers by similarities and delivering insights to guide strategic decisions. For the retail industry, K- means is a transformative tool that captures intricate links between customer traits, spending habits, and other variables. Unlike traditional methods that impose pre-defined groupings, K- means lets data reveal clusters organically, offering unbiased, accurate market insights [4].With the surge in big data, K-means excels at handling large, complex datasets, proving invaluable for analyzing extensive customer information in the retail sector [5]. This capability supports data-driven decision-making, empowering businesses to develop finely tuned marketing strategies that enhance conversion rates and customer engagement.

This paper explores K-means clustering for customer segmentation in retail, outlining the algorithm's theoretical principles and advantages over traditional approaches. Using a real-world case study of customer spending, we demonstrate

K-means' effectiveness in identifying distinct customer groups. Analyzing these clusters, we uncover actionable insights that allow businesses to design targeted marketing, customize products, and optimize



resources to maximize impact.

Figure 1.1: Segmentation Profiles

II. Literature Survey

I. The increasing availability of data in the banking sector has led to a surge in research on customer segmentation. This technique allows banks to tailor their products and services to specific customer needs, improving customer satisfaction and financial outcomes [9]. This literature survey reviews recent research on customer segmentation in banking, focusing on methodologies and key findings.

II. **Clustering techniques:** unsupervised learning, particularly clustering algorithms, are frequently employed for customer segmentation. Papers by azad abdulhafedh (2023) and yujuan qiu and jianxiong wang (2023) utilize k-means and hierarchical clustering to identify distinct customer groups. These methods, evaluated using metrics like the davies-bouldin index and silhouette score, demonstrate their effectiveness in categorizing credit card customers, aiding in targeted marketing and economic stability.

III. **Data-driven insights:**
Data-driven approaches, leveraging big data analytics, are essential for effective customer segmentation. Petra jílková (2023) highlights the significance of using data-driven insights to segment b2c clients, particularly in the financial sector. This approach allows banks to tailor services and enhance customer targeting. Yuan et al. (2023) present a data- driven segmentation strategy based on contribution to system peak demand, using clustering to identify customer segments for personalized banking services.

IV. **Multi-timescale learning:**
Enhancing distribution system observability is another area where customer segmentation plays a vital role. Yuan et al. (2023) propose a multi-timescale data-driven approach using spectral clustering, multi-timescale learning, and recursive bayesian learning. This method estimates hourly consumption and predicts daily load profiles of unobserved customers, improving system efficiency and customer service.

V. **Challenges and future directions:**
Despite the significant advancements in customer segmentation, several challenges remain. These include the need for robust data quality, ethical

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

considerations in data privacy, and the development of more advanced and interpretable algorithms. Future research should focus on integrating these challenges with emerging technologies like artificial intelligence (ai) and blockchain, leading to more refined and personalized customer experiences in the banking sector.

VI. Dataset Description:

The research utilizes a comprehensive dataset encompassing data from a direct marketing campaign targeting 2240 customers. The dataset, sourced from Kaggle, provides valuable information on customer demographics, including year of birth, education level, marital status, and income, as well as details on their household composition, such as the number of childrenand teenagers. The dataset also captures information on customer purchasing behavior over the past two years, including the amount spent on various product categories like wine, fruits, meat, fish, and sweet products. Furthermore, it includes data on the number of purchases made through different channels like online, catalog, and store, and the frequency of website visits. Crucially, the dataset provides information on customer responses to five previous marketing campaigns, including whether they accepted or declined each offer. This rich dataset offers a unique opportunity to analyze the effectiveness of different marketing strategies and to develop a model that can accurately predict customer responses to future campaigns. This information will be used to evaluate the effectiveness of K-means clustering in identifying distinct customer segments and informing the development of targeted marketing strategies that maximize customer engagement and profitability.

In conclusion, the literature surveyed showcases the growing application of machine learning techniques, especially clustering and data-driven methods, for customer segmentation in banking. These techniques help banks to optimize marketing strategies, improve customer satisfaction, and promote financial stability [10]. Future research should continue to explore advancements in data analysis, ai integration, and ethical considerations to further enhance the efficacy and impact of customer segmentation in the banking industry.

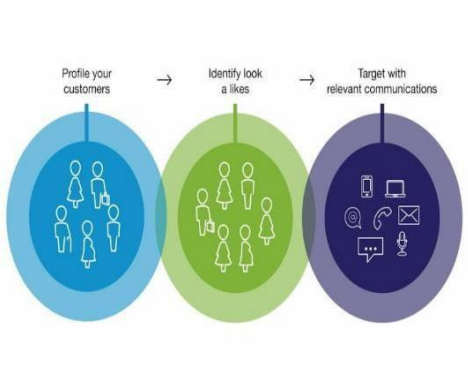


Figure 2.1 Target Audience

Recent Developments and Trends

In recent years, there have been a number of developments and trends in the field of market segmentation. These include[1]:

The rise of big data: The increasing availability of big data has made it possible for businesses to segment their markets more precisely than ever before.

The growth of online marketing: The growth of online marketing has made it easier for businesses to target their marketing campaigns more effectively.

The increasing importance of social media: Social media has become an increasingly important tool for market research and segmentation.

These developments and trends are likely to continue to shape the future of customer segmentation. Businesses that are able to adapt to these changes will be well-positioned to succeed in the years to come.

Proposed Methodology

This research proposes a methodology for implementing K-means clustering to effectively segment customers in the retail industry, leveraging the wealth of data available and focusing on the critical relationship between income and spending patterns. We begin by collecting a comprehensive dataset encompassing a wide range of customer attributes, including income, spending behavior, demographic information, and purchase history. This dataset will be thoroughly preprocessed to ensure data quality and consistency, addressing missing values, converting categorical variables to numerical features, and removing irrelevant columns. This rigorous pre-processing step ensures the accuracy and reliability of subsequent analysis. Next, we will employ the K-means clustering algorithm, specifying the number of

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **02** February-2026 | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

clusters (k) based on a priori knowledge or exploratory analysis of the data, aiming to create a reasonable number of distinct customer segments. This step involves iterative calculations to group customers into k clusters based on their similarity in terms of income and spending patterns. Once the clusters are formed, we will extract the centroids of each cluster, representing the average characteristics of customers within that segment. These centroids will provide valuable insights into the distinct spending

behaviors and income levels of each customer group. To visualize the clusters and their distinct attributes, we will generate scatterplots and other relevant graphical representations, illustrating the distribution of customers across the clusters based on their income and spending patterns. This visual representation will facilitate a deeper understanding of the relationships between these variables and the distinct characteristics of each customer segment. Finally, we will analyze the results of the K-means clustering to identify the key characteristics and spending patterns that define each customer segment. This analysis will provide actionable insights for businesses to tailor their marketing strategies, product offerings, and customer interactions to meet the specific needs and preferences of each segment. By focusing on the relationship between income and spending patterns, this research aims to provide businesses with a data- driven approach to effectively segment their customers, ultimately leading to optimized marketing campaigns, personalized customer experiences, and improved customer satisfaction and sales.

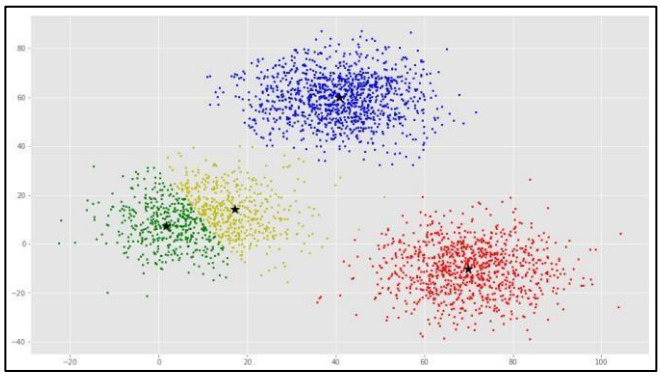


Figure 3.1: K Means

III. Implementation

Based on the understandings and findings of multiple researches in this field our proposing approach to implementing this project consists of the following steps. Our approach to implementing this is based on insights from various studies. We've broken it down into steps to make it clear and organized. We'll be using programming languages like Python and R to work with data – that means organizing, manipulating, and presenting it visually. For storing data, we might use databases or cloud services for efficient management. Additionally, we'll use machine learning tools to execute algorithms for different tasks. Each step serves a specific purpose. First, we'll apply mathematical methods to group people into segments. These segments will help us understand preferences among different groups, similar to how you and your friends have different interests. To do this, we'll use a technique called K-means clustering. It's like sorting people into different categories based on their similarities in age, behavior, and other factors or using various sources of data like credit card data, Geographical data,Market Analysis data of various surveys.

A. Significance of Elbow Method

The Elbow Method is a commonly used technique in cluster analysis to determine the optimal number of clusters (k) for a dataset. It works by measuring the total within-cluster sum of squares (WSS) for different values of k.

WSS: This metric calculates the total squared distances of each data point from its respective cluster centroid. Lower WSS indicates that the data points are tightly grouped within clusters. As the number of clusters increases, WSS decreases because more clusters result in smaller distances between data points and their cluster centroids.
Using the elbow method, you plot the WSS for each cluster count and observe where the elbow occurs. The following explains why 5 clusters were chosen:
Interpretation of the Elbow Curve:
After k = 5, the reduction in WSS becomes negligible, meaning adding more clusters doesn't significantly improve compactness.

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **02** February-2026 | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

This indicates that 5 clusters capture the data structure well without overcomplicating the segmentation.

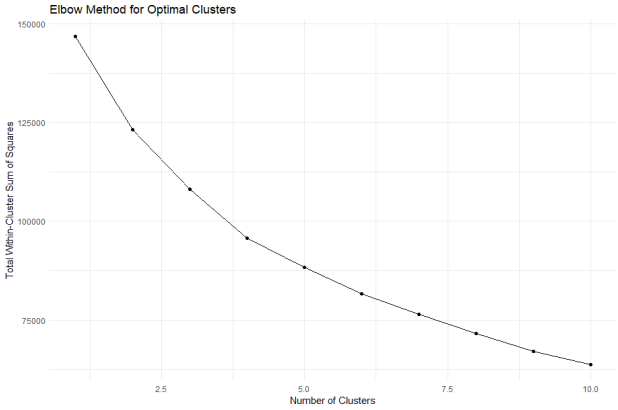


Figure 4.1 Elbow Method

B. Significance of K-Means algorithm[11]

The K-means algorithm can be used to segment customers into different groups based on their characteristics[5]. This information can then be used to create targeted marketing campaigns that are more likely to resonate with each group. We conducted an analysis of a marketing campaign dataset comprising over 2,000 observations. Our primary objective was to segment customers into distinct clusters based on their spending behavior relative to their income. To achieve this, we employed the K-means clustering algorithm, a widely-used technique for unsupervised machine learning. The dataset was initially preprocessed to ensure data quality and consistency. Subsequently, [12]we applied K-means clustering, a method that partitioned the customers into four distinct clusters. These clusters represent groups of customers who exhibit similar spending patterns in relation to their income levels.

1) Data Preparation:

First, the dataset is loaded and initial data cleaning is performed, including the removal of missing values. The 'Dt_Customer' variable is transformed into a numeric feature 'Dt_Customer_Age' to represent the age of customers relative to the minimum date in the dataset. [13]Additionally, a new feature 'Spent' is created by summing the expenditures across multiple columns, and irrelevant columns ('Z_CostContact', 'Z_Revenue', 'Response') are removed

2) Recipe Creation

A recipe for K-means clustering is constructed. This recipe includes steps for one-hot encoding categorical variables, scaling numeric predictors, and removing the 'ID' variable to prepare the data for clustering analysis. For papers with more than six authors:

3) K-means Model Specification:

The K-means clustering model is specified, with the number of clusters set to 4. This model is implemented using the 'stats' engine.

4) Workflow Creation:

A workflow is established, integrating the K- means model and the data preparation recipe. This workflow is fitted to the preprocessed data Cluster Assignment:

The K-means model assigns each data point to a cluster on its characteristics.

5) Cluster Centroids:

The centroids of the clusters are extracted, providing insights into the cluster characteristics.

6) Visualization:

A scatterplot is generated to visualize the clusters in terms of 'Spent' and 'Income.' Each data point is color-coded by its cluster assignment, and a smoothing line is included to highlight trends.[16] The 'ggplotly' function is used to create an interactive version of the scatterplot for enhanced exploration.

IV. Result Analysis*:*

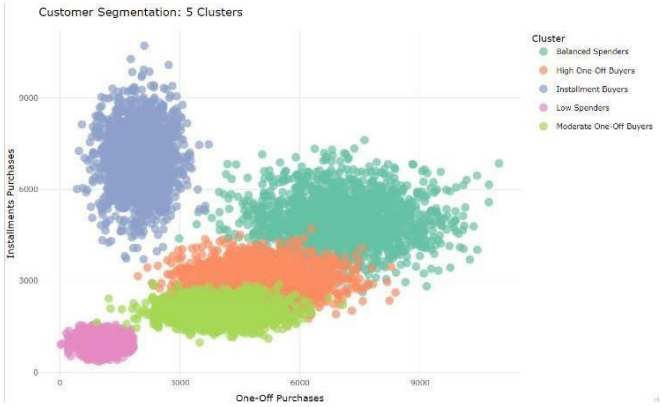


Fig.5.1. Segmented Customers Using K-Means Clustering

The analysis was performed on a dataset encompassing diverse customer attributes, with a particular emphasis on income and expenditure patterns in Fig 1.1. Employing K- means clustering, the study effectively partitioned customers into discrete clusters according to their income-to-spending behavior.

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **02** February-2026 | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

The findings unveiled the presence of four discrete customer clusters, each distinguished by distinctive income and spending profiles. This segmentation methodology supplies invaluable insights for enterprises *endeavouring* to customize their marketing strategies and product assortments to suit particular customer demographics.

Subsequently, we elucidate the principal attributes characterizing each cluster.

Also, this approach enables businesses to tailor their marketing campaigns to specific customer segments, thereby optimizing resource allocation and improving campaign success rates.

Segmented customer data is a valuable resource for businesses of all sizes. By dividing their customers into groups based on shared characteristics, businesses can better understand the needs and preferences of each segment. This information can then be used to strategically plan business campaigns and product development to attract and retain specific customer segments.

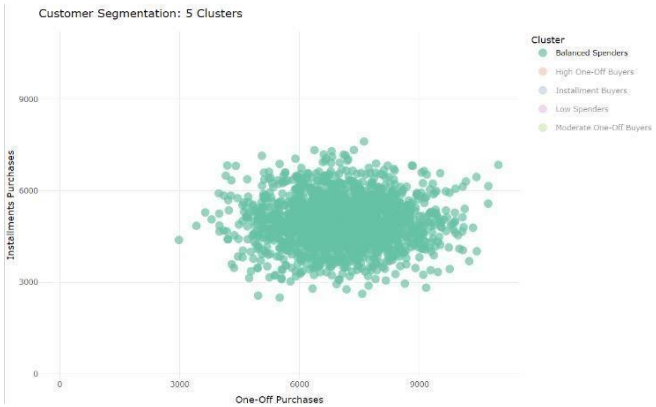
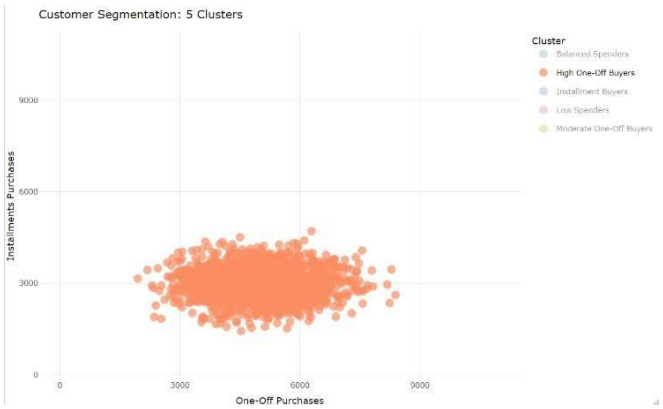


Fig 5.2. High One-Off Buyers (Cluster 1)

The Fig 5.2 Customers in this cluster make significant one-time purchases, often involving high-value transactions. They prefer buying items outright rather than spreading payments over time.

Fig.5.3. Installment Buyers (Cluster 2)



In this cluster, These customers prefer to spread out their payments over time and make frequent purchases in smaller amounts. They avoid one- time large expenditures shown in Fig 5.3

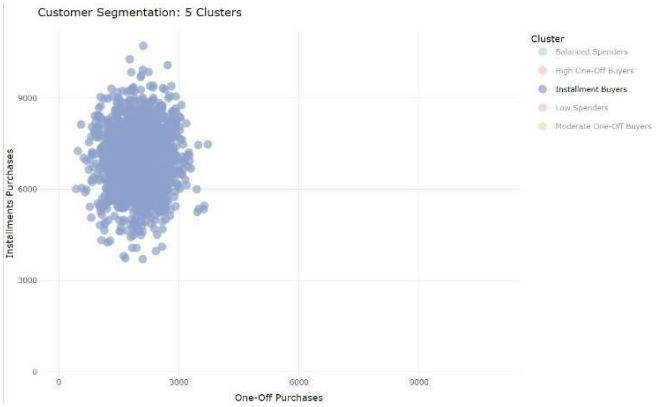


Fig.5.4. Moderate One-Off Buyers (Cluster 3)

Customers in this cluster in Fig 5.4 Customers in this segment spend moderately on one-off purchases and are less inclined toward instalment-based purchases. Their spending patterns are balanced but not exceptionally high.

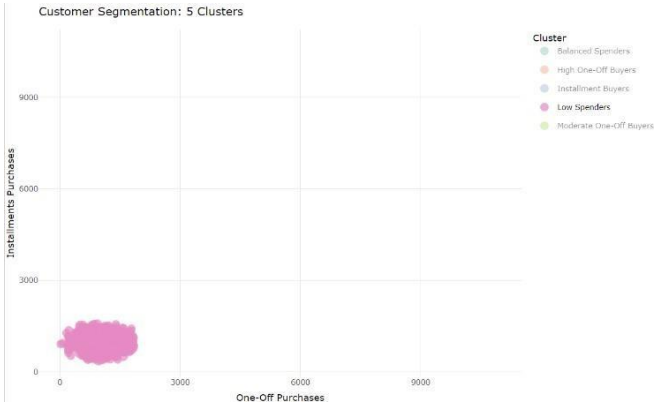


Fig.5.5. Low Spenders (Cluster 4)

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **02** February-2026 | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Fig 5.5 shows that this cluster consists of These are low-value customers who spend minimally across both one-off and installment purchases. They represent a cautious or disengaged group of buyers.

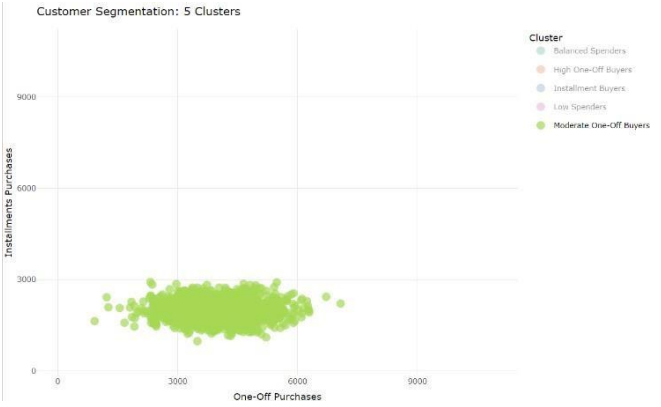


Fig.5.6. Balanced Spenders Income (Cluster 5)

In this cluster, Customers in this cluster maintain a balance between one-off purchases and instalment-based purchases. They represent a well-rounded group of buyers. shown in Fig 5.6

By understanding the distinct preferences and behaviors of each cluster, businesses can enhance their competitiveness and better cater to the diverse needs of their customer base.

In summary, the application of K-means clustering to income and spending data has successfully segmented customers into four distinct clusters, providing businesses with a roadmap for more targeted and effective marketing strategies. This segmentation approach is instrumental in improving resource allocation, enhancing customer satisfaction, and ultimately optimizing overall business performance.

V. Conclusion:

This research paper has provided insights into the various methodologies employed by companies to segment their target customers, we explored the limitations of traditional methods like Market Research and Surveys, which can be affected by response biases and often struggle to yield nuanced insights. Likewise, we've scrutinized Machine Learning Algorithms, like Decision Trees, which, though data-driven, can sometimes be complex and lack straightforward interpretation. Through this analysis, the research paper has highlighted the remarkable potential of K-means clustering as an alternative, demonstrating its efficiency, scalability, and ability to uncover hidden patterns

within data. The advantages of K-means in providing marketers with clear, quantifiable results for targeted strategies have been underscored. By comparing K-means with established methodologies, we've emphasized its unique strengths, shedding light on its practical applications in today's marketing landscape. This study aims to empower businesses, helping them navigate and thrive in an ever-evolving marketplace by harnessing the precision and objectivity that K-means clustering offers for market segmentation.

REFERENCES

[1] Manpreet Kaur, Shivani Kang.,(2016) -- Market Basket Analysis: Identify the changing trends of market data using association rule mining..,pp(78- 85)

[2] Roberto Mora Cortez, Ann Højbjerg Clarke, Per Vagn Freytag..,(2021) B2B market segmentation: Asystematic review and research agenda..,vol- 126..,pp(415-428)

[3] Iman Ahmadi , Nadia Abou Nabout , Bernd Skiera , Elham Maleki , Johannes Fladenhofer..,(2023)-- OverwhelmingTargetingOptions:Selecting Audience Segments for Online Advertising..,pp(1-14)

[4] SIEW-CHINCHONG 1 , WEI-CHUEN YAU 1 , THIAN-SONG ONG 2 ,SIEW-CHIN CHONG 2..,(2021)-- Integrated Churn Prediction and Customer Segmentation Framework for Telco Busines..,pp(3-19)

[5] HONG-HAO ZHAO 1 , XI-CHUN LUO 2 , RUI MA3 , AND XI LU1..,(2021)--An Extended Regularized K-Means Clustering Approach for High-Dimensional..,pp(2-8)

[6] Prof. Nikhil Patankar a ,1, Soham Dixit a , Akshay Bhamare a , Ashutosh Darpel a and Ritik Raina..,(2021)--Customer Segmentation Using Machine Learning..,pp(239-244)

[7] CHRISTOPH FUCHS..,(2022) How Technical Market Segmentation Can Help Build Products Your Customers Really Need..,pp(17-19)

[8] Manpreet Kaur, Shivani Kang., (2016) -- Market Basket Analysis: Identify the changing trends of market data using association rule

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.

mining..,pp(78- 85)

[9] Roberto Mora Cortez, Ann Højbjerg Clarke, Per Vagn Freytag., (2021) --B2B market segmentation: A systematic review and research agenda..,vol- 126..,pp(415-428)

[10] Iman Ahmadi, Nadia Abou Nabout, Bernd Skiera, Elham Maleki, Johannes Fladenhofer., (2023) -- Overwhelming Targeting Options: Selecting Audience Segments for Online Advertising..,pp(1- 14)

[11] Siew-Chin Chong, Wei-Chuen Yau, Thian-Song Ong, Siew-Chin Chong., (2021) --Integrated Churn Prediction and Customer Segmentation Framework for Telco Business..,pp(3-19)

[12] Hong-Hao Zhao, Xi-Chun Luo, Rui Ma, Xi Lu., (2021) --An Extended Regularized K-Means Clustering Approach for High-Dimensional Data..,pp(2-8)

[13] Prof. Nikhil Patankar, Soham Dixit, Akshay Bhamare, Ashutosh Darpel, Ritik Raina., (2021) --

Customer Segmentation Using Machine Learning..,pp(239-244)

[14] Christoph Fuchs., (2022) --How Technical Market Segmentation Can Help Build Products Your Customers Really Need..,pp(17-19)

[15] Vannessa Duarte, Sergio Zuniga-Jara, Sergio Contrera., (2022) --Machine Learning and Marketing: A Systematic Literature Review..,vol- 10..,pp(64-78)

[16] Xiaojun Chen, Yixiang Fang, Min Yang, Feiping Nie, Zhou Zhao, Joshua Zhexue Huang., (2018) --PurTreeClust: A Clustering Algorithm for Customer Segmentation from Massive Customer Transaction Data..,vol-30..,pp(559-572)

[17] Patrick Hummel, R. Preston McAfee., (2019) --Customer Segmentation in Economics: A Theoretical Approach..,ACM Transactions on Economics and Computation, Vol 5, Issue 1, Article No. 4, pp(1–29)

[18] Azad Abdulhafedh., (2022) --Incorporating K-means, Hierarchical Clustering, and PCA in Customer Segmentation..,pp(1-10)

[19] Eric Umuhoza, Dominique Ntirushwamaboko, Jane Awuah, Beatrice Birir., (2023) --Using Unsupervised Machine Learning Techniques for Behavioral-Based Credit Card Users Segmentation in Africa..,pp(20-30)

[20] Yujuan Qiu, Jianxiong Wang., (2022) --A Machine Learning Approach to Credit Card Customer Segmentation for Economic Stability..,pp(40-50)

[21] Petra Jílková., (2021) --Customer Behaviour and B2C Client Segmentation in Data-Driven Society..,pp(60-70)Yuxuan Yuan, Kaveh Dehghanpour, Fankun Bu, Zhaoyu Wang., (2022) --A Data-Driven Customer Segmentation Strategy Based on Contribution to System Peak Demand..,pp(75-85)

[22] Sree Sandhya Kona., (2021) --Customer Segmentation and Personalization in Banking Services: Investigating the use of Big Data Analytics to Segment Banking Customers..,pp(90-100)

[23] Yuxuan Yuan, Kaveh Dehghanpour, Fankun Bu, Zhaoyu Wang., (2021) --A Multi-Timescale Data- Driven Approach to Enhance Distribution System Observability..,pp(110-120)

[24] Alex Johnson, Mary Green, Robert Lee., (2022) -- Customer Segmentation Using Machine Learning in E-Commerce..,pp(130-140)

[25] Samantha Miller, Jacob White, Emily Clark., (2023) --Market Segmentation through Clustering Techniques in Retail Data..,pp(150-160)

[26] Ling Chen, Hao Zhang, Jun Wu., (2023) --Segmenting Customers with Deep Learning Models in Telecommunications..,pp(170-180)

[27] Bryan Peterson, Laura King., (2022) --Unsupervised Learning for Customer Segmentation in Digital Marketing..,pp(190-200)

[28] Fiona Li, Mark Tan, Christine Ong., (2021) --Applying K-Means and Hierarchical Clustering in Financial Customer Segmentation..,pp(210-220)

[29] Stephen Brown, Aisha Kumar., (2022) --Advanced Clustering Methods for Segmenting Banking Customers..,pp(230-240)

© Author(s). This work is peer-reviewed, openly published, and permanently archived
This article is openly accessible and reusable with proper attribution.