

Analysis of consumer characteristics on retail business with clustering analysis method and association rule for selling improvement strategy recommendations

Muhammad Rafly Qowi Baihaqie¹, Annisa Uswatun Khasanah^{1*}

¹Department of Industrial Engineering, Universitas Islam Indonesia, Jalan Kaliurang km 14,5, Yogyakarta

*Corresponding Author: annisa.uswatun@uii.ac.id

Article history

Received: 10 December 2023

Revised: 9 June 2024

Accepted: 14 June 2024

Published: 30 June 2024

Keywords:

Retail

Customer segmentation

Customer buying pattern

Clustering analysis

Association rule

ABSTRACT

In the highly competitive retail industry, companies must continually innovate and develop unique business strategies to enhance their sales performance. The ABC Store, a mini market in Yogyakarta, has experienced fluctuating sales over the past year, failing to meet its targets. This study aims to analyze consumer purchasing behavior at the ABC Store and provide strategic recommendations to boost sales. The data analyzed in this study comprises three months of transaction records. The methods used include Association Rule - Market Basket Analysis (AR-MBA) with the FP-Growth algorithm and Clustering Analysis with K-Means. The clustering analysis identified four distinct customer segments: Mid-Morning Moderates, Diverse Afternoon Buyers, Evening Moderates, and High-Value Customers. Cluster 2, comprising Diverse Afternoon Buyers, was selected for AR analysis due to its relatively high transaction value and the variety of products purchased, indicating its potential to evolve into a High-Value Customers cluster. The analysis yielded 104 rules. The findings can inform marketing strategies to increase sales, including product bundling and customer loyalty programs such as a point system.

DOI:

<https://doi.org/10.31315/opsi.v17i1.11411>

This is an open access article under the [CC-BY](https://creativecommons.org/licenses/by/4.0/) license.



1. INTRODUCTION

The retail industry in Indonesia continues to grow. According to data from Euromonitor, total retail outlets in Indonesia reached 3.98 million by 2022 [1]. It makes the retail industry in Indonesia more competitive and forces companies to continue to improve performance and maintain superiority [2]. However, in the context of such increasingly positive retail growth, many retail stores are experiencing difficulties and even failures. According to Budihardjo Iduansjah, General Chairman of Himpunan Peritel dan Penyewa Pusat Perbelanjaan Indonesia (Hippindo), one of the factors causing the change in consumer behavior in shopping. Surveys show that people prefer to shop at small retail stores to save time [3]. As proof of this phenomenon is PT Hero Supermarket Tbk (HERO) which officially in July 2019 closed 6 GIANT stores. The closure of these large retail stores is due to the stringent offline retail competition and changes in the company's business strategy [4]. In contrast, PT Alfaria Trijaya Tbk. (AMRT), which owns the Alfamart, Alfamidi, and Lawson brands, grew to 9.3% or as much as 1,283 stores in 2018. Also, PT Indomarco Prismatama, which holds the Indomaret brand, also experienced growth of 6.7% or added 1.031 stores in 2018 to 16.366 stores [4, 5].

The ABC Store, a retail chain in Jogja with 10 branches, faces similar challenges in a competitive market that requires constant innovation. Despite various marketing strategies, the store has experienced poor sales trends, particularly in 2022. Sales data reveal that targets are met only during specific months, such as Ramadan and Idul Fitri. This inconsistency poses a significant concern for the management, who need to ensure steady sales throughout the year. Efforts to boost sales, like offering discounts on specific brands, have not yielded significant results. Currently, the store's discount strategy is not specialized and relies solely on the price reductions provided by the brands themselves.

The forms of innovation strategy for the retail industry can be very different. A business without good marketing will not be attractive to customers, which in the end can affect the sales and sustainability of the retail industry. Business will not last long without good and sustainable marketing. A proper and strong marketing strategy will separate between a successful retail business and a business that can't attract many customers [6].

In designing an effective marketing strategy, data is needed as a reference. Data is a powerful tool in making corporate policies or strategies. Nowadays, it's called the age of big data, where in a short time, data increases very rapidly and with high variation. The retail sector is one of the industries that do big data analytics or big data, which is very promising and something different for retailers. The IBM Global Business Services report found the fact that 62% of retailers who use information analysis through big data create a competitive advantage for their companies compared to 63% of other industry respondents. [7].

The ABC Store holds a substantial amount of data in its database. Among this wealth of information, the transaction data remains underutilized, despite its potential to offer valuable insights into customer behavior, sales trends, product preferences, and business needs. Leveraging transaction data is crucial for enhancing business efficiency and profitability. This research aims to develop an effective marketing strategy for the ABC Store by analyzing the available data. Specifically, the data will be used to segment customers and identify purchasing patterns.

By conducting segmentation analysis using sales transaction data, an understanding of the different customer characteristics within several groups can be established. Customers will be grouped according to their purchasing behavior so that it can be easier to implement the right targeted marketing strategy [5]. One method that can be used to group consumers or segmentation is the clustering method. Clustering is a method to group data based on the similarity of characteristics between one data and another [8]. One of the popular algorithms in the clustering method is the k-means. This algorithm works with stages, i.e. determines the number of clusters desired, performs data allocation to a given cluster number, determines centroid values on each cluster, calculates the closest distance between each data point using the Euclidean formula, and displays the calculations based on the distance [9]. The K-means algorithm has the advantage of being able to group big data very quickly [10]. The K-means algorithm has an advantage in time efficiency when iterating with large amounts of data compared to the K-medoids method. Thus, K-Means can provide a significant advantage in faster and more efficient data grouping. This study aims to use the K-means algorithm instead of other clustering methods, such as hierarchical clustering, due to the large dataset involved. Hierarchical clustering presents significant challenges in visualizing the hierarchy when applied to large datasets. K-means is widely applied by previous authors in customer segmentation such as Sarkar et al. [11], Marisa et al. [12], Li et al. [13], and Siagian et al [14].

In addition to using the Clustering method, another method used in this research is the Market Basket Analysis (MBA) method with the Association Rule Mining. Market basket Analysis is a data analysis technique used to find patterns of product purchase using sales transaction data. The primary objective of the MBA in Marketing is to provide information to the store owner to understand the purchasing behavior of the buyer, which can help in making the right decisions [15]. One of the algorithms that can be used in the Market Basket Analysis (MBA) method with the Association Rule Mining is the FP-Growth. FP-Growth is an algorithm used to identify frequent item sets in data sets without using candidate generation [16]. The FP-Growth algorithm uses the tree data structure in the frequent itemset search process, thus being able to perform searches more efficiently and quickly compared to the a priori method [17]. Using this approach, the FP-Growth algorithm can avoid a process called candidate generation, which can be time-consuming and consume enormous computing resources. As a result, the FP-growth algorithm has become one of the popular choices for analyzing association patterns in big data. Erwin [18] mentions that the apriori algorithm has shortcomings in terms of data processing time and large memory allocation, this is due to repeated data scanning. On the other

hand, the FP-Growth algorithm uses the data structure of the F P-Tree that can compress the transaction data with the same item, thus making more efficient memory usage and the frequent itemset search process faster. Some authors also conduct studies using AR algorithms such as Dwiputra et al. [19], Wibowo et al. [20], and Putri et al. [21]. Some studies in different scope also integrated Clustering analysis and Association Rule, such as Hao et al. [22], Kusak et al. [23], Dol et al. [24], and Rahman et al. [25]. Hao et al [22] used the association rule and improved K-means algorithm to overcome the problems of time-consuming, low-precision, and redundant rules in association rule mining of big data. Kusak et al. [23] used the association rule and K-means algorithm to interpret pre-event landslide areas and landslide inventory. While Dol et al. [24] and Rahman et al. [25] implemented the association rule and K-means algorithm in the educational field.

Considering the challenges faced by the ABC Store, this research aims to propose an effective sales enhancement strategy to help the store achieve its sales targets. The strategy will be developed using Association Rule-Market Basket Analysis (AR-MBA) and Clustering methods. The AR-MBA method will identify purchasing patterns among items in customer transactions, enabling the store to make relevant sales recommendations. Meanwhile, the clustering method will segment customers based on transaction data, assisting in classifying distribution patterns and uncovering interesting correlations between data attributes. By applying both approaches, the ABC Store can devise more targeted sales strategies and improve personalization in product recommendations for customers.

2. METHODS

The research was conducted at the ABC Store, a retailer in Yogyakarta, using three months of transaction data from November 2022 to January 2023. Following data collection, the study proceeded through four stages: data preprocessing, customer segmentation analysis using clustering methods, identification of customer purchasing patterns with Association Rule-Market Basket Analysis (AR-MBA), and the formulation of appropriate marketing strategies. All stages of the study were executed using the Python programming language.

During data preprocessing, the data for cluster analysis and AR-MBA are separated and preprocessed individually to ensure optimal quality and to tailor it to the specific requirements of clustering and AR-MBA analysis. Data cleaning addresses noise, missing, and incomplete data. For AR-MBA, transactions with only one purchase item are eliminated to focus on transactions with multiple items, which are more relevant for identifying association patterns. Data transformation customizes the data to suit the needs of clustering and AR-MBA analysis. For clustering, an example of data transformation includes adding a "time" attribute column that indicates the time of each transaction. For AR-MBA, transactions with the same ID are combined to ensure each transaction ID represents a unique transaction. This step also involves converting transaction data into a Boolean format, where each item is represented as True or False based on its presence in the transaction, allowing for efficient analysis of association patterns. In the data selection stage, relevant variables or attributes for clustering analysis are chosen. This involves selecting a subset of data with important and relevant information for the clustering process. The final stage is data normalization, which aims to equalize the data scale so that no single large-scale data point dominates the clustering process.

Once the data has undergone preprocessing, it is ready for analysis. The first phase is cluster analysis, which aims to identify the characteristics of different customer segments. The attributes used for cluster analysis include customer transaction times, the quantity of goods purchased, and total transaction amounts, all of which are recorded in the consumer transaction data. The K-means clustering algorithm is employed for this analysis. To determine the optimal number of clusters, both the elbow method and the silhouette method are used. Once the optimal number of clusters is established, the K-means algorithm partitions the data into k clusters by identifying the centroids that represent each data group. The data points are then assigned to clusters based on the Euclidean distance between each point and the nearest centroid.

Once the cluster analysis is completed and customer segment profiles are identified, the next step is to analyze purchasing patterns within the target customer cluster. This involves using the FP-Growth algorithm to identify item sets that frequently appear together in transactions. After generating the association rules, researchers evaluate and select the most relevant and meaningful rules using parameters such as support, confidence, and lift. Additionally, business factors and analytical needs are considered to choose the most suitable rules. The objective of this phase is to obtain association rules with high support and confidence values that are relevant to the analysis.

The final phase of the research is to formulate sales enhancement strategies based on the results from the clustering analysis and AR-MBA. These strategies are developed because both methods provide valuable insights into customer behavior and product purchasing patterns in the store. By leveraging the information from the clustering results, the store can create marketing strategies that are more targeted and tailored to the preferences of each customer group. Furthermore, by applying the association rules derived from the AR-MBA method, the store can create attractive product bundling packages and optimize the promotion and sale of products that have strong associative relationships.

3. RESULTS

The dataset used in this study, after preprocessing, consisted of 31,735 transactions, including information on transaction times, variations of purchased goods, total transaction amounts, and types of goods purchased.

3.1. Clustering Analysis

The attributes used in this cluster analysis are the number of product variations purchased by the customer, the total transaction amount, and the time of customer does the transaction at the store [26]. In clustering analysis, the "time" attribute is utilized to identify the ideal hours for promotions or marketing activities. This information helps researchers determine the optimal times to attract customer interest and enhance the effectiveness of product promotions. The "product variation" attribute is selected to examine whether there are differences in the quantity and variety of goods purchased by customers at different times. Researchers aim to identify trends in customer preferences related to the number of product variations bought at specific times. Additionally, the "total transaction" attribute provides insight into how much money customers spend at different hours and on various items. This helps understand customer spending patterns and the value of particular shopping carts, which is crucial for assessing potential revenue and the contribution of transactions to the business.

The study employed the elbow method and silhouette score to determine the optimal number of clusters for the clustering analysis. The elbow method produced a graphical plot with cluster numbers ranging from 2 to 10. In Figure 1, inflection points appeared at clusters 3 and 4. To ensure accurate cluster selection, the silhouette score method was also applied. During this phase, the silhouette scores were plotted, and it was noted that cluster 4 had the highest silhouette score compared to clusters 3 and 4, as shown in Figure 2. This suggests that cluster 4 had a better level of data cohesion and separation than the other clusters. Therefore, four clusters were used for further analysis of customer behaviour, purchasing preferences, and spending patterns in the store's business context.

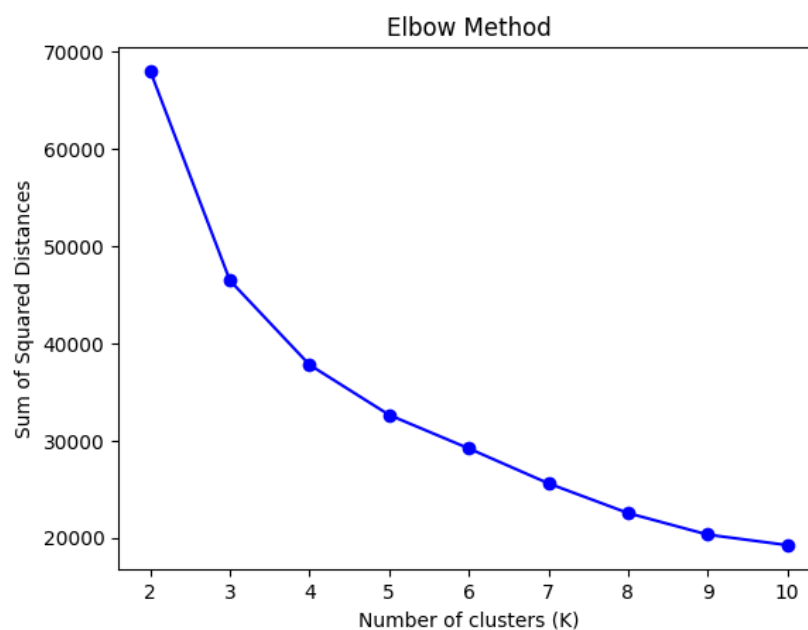


Figure 1. Elbow method graph

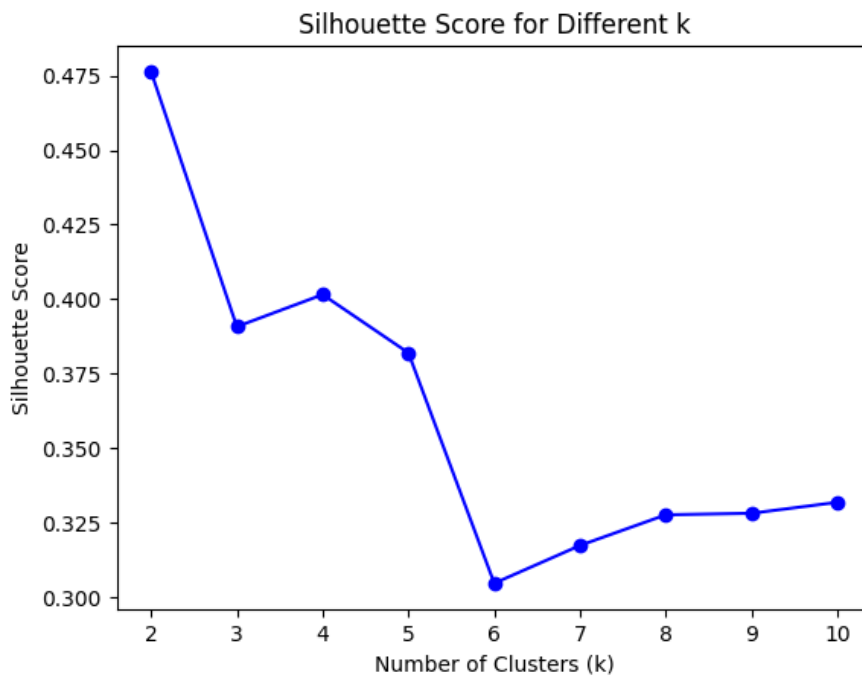


Figure 2. Silhouette Score graph

The clustering process using the K-Means algorithm was performed with VSCode software and the Python 3.11.4 programming language, yielding the results shown in Table 1.

Table 1. Clustering analysis result

Cluster	Cluster member	Percentage (%)	Transaction time	Product variation	Total transaction (Rp)
1	12825	40,4	10.33	1.58	19866.89
2	4191	13,2	15.88	4.83	49271.20
3	14172	44,7	18.25	1.56	19591.12
4	547	1,7	13.87	4.73	177475.84

On cluster 1, the known average transaction time was 10.33 or around 10 am., with data spread between 8 and 12. This shows that most transactions in this cluster were between 8 am and 12 pm. For the Product Variation attribute, the average is 1.6, with a data spread between 1 to 2 product variations. This result shows that this cluster reflects customers who tend to buy products in small quantities. For the Total Transaction Variable, the mean transaction value is Rp. 19.866, - with the data range from Rp. 8.925, - to Rp. 26.540, -. It indicates that customers in these clusters have transaction values that are in the medium or medium range. These clusters can be called "Mid-Morning Moderates", because they have transaction values in a medium or medium range. They tend to buy the same product or have little variation and make transactions from morning to noon. "Mid-Morning" describes the dominant time when customers in this cluster make transactions. "Moderates" emphasizes stable and balanced purchasing rates. The researchers named the cluster based on its defining characteristics.

On cluster 2, it is known that the average time of arrival was 15.9 or around 3 pm to 4 pm., with data spread between 1 pm and 7 pm. It shows that most of the transactions in this cluster are done from 1 pm to 7 pm. For the Product Variation attribute, the average value obtained is 4.8, with a data spread between 4 to 5 product variations. This cluster shows customers who are inclined to buy different variations of the product. For the Total Transaction attribute, the mean value of the transaction is Rp. 49.271 with the data range from Rp. 31.000 to Rp. 63.434. This suggests that customers in these clusters have transaction values that are in a fairly high range. They tend to buy a variety of products and make transactions from day to evening. "Diverse"

describes the tendency of customers in this cluster to try the range of products offered. "Afternoon Buyers" emphasizes the dominant time when customers make transactions.

On cluster 3, it is known that on the Transaction Time attribute, the average obtained is 18.3, with a data spread between 17 and 20. This shows that most transactions in this cluster are done between 5 pm and 8 pm. For the Product Variation attribute, the average obtained is 1.6, with a data spread between 1 to 2 product variations. This cluster is similar to cluster 1, where customers tend to buy products with little variation. For the Total Transaction attribute, the mean transaction value is Rp. 19.591 with the data range from Rp. 9.500 to Rp. 26.400. This suggests that customers in these clusters tend to make transactions with values that are in the medium or medium range, not too low but not too high. These clusters are a group of customers who do transactions from afternoon to night, choose products with limited variation, and are oriented towards stable and consistent transactions. These clusters can be called "Evening Moderates" because they reflect customers who have transaction values in a medium or medium range. They tend to buy products with little variation and make transactions in the afternoon to the night. "Evening" describes the dominant time when customers in this cluster do the transaction. "Moderates" emphasizes the customer's tendency to trade at a stable and balanced rate.

As for cluster 4, for the Transaction Time attribute, the average obtained is 13.9, with a data spread between 10 and 18. This shows that most transactions in this cluster are done between 10 a.m. and 6 p.m. For the Product Variation attribute, the mean obtaining is 4.7, with data spreads between 1 to 7 product variations. This cluster shows customers who are inclined to buy various variations of products with an average transaction value of Rs. 177.475.-. These clusters reflect customers who perform several transactions at a very high value, far above the average transaction of their other cluster so they have a significant contribution to the store's revenue. These clusters can be called "High-Value Customers", this cluster can be referred to as "high-value Customers" because they reflect customers who have very high transaction value. This reference describes that this cluster represents a very valuable customer group and can be a major focus in marketing strategies aimed at increasing customer loyalty with high transaction value.

Of the four clusters, the best cluster for improvement and recommendation is Cluster 2, the "Diverse Afternoon Buyers". This cluster has transaction values that are in a fairly high range and customers tend to buy a variety of product variations. It's hoped that customers in cluster 2 can be customers in cluster 4.

3.2. Association Rule Market Basket Analysis

Customer purchase pattern analysis is performed for transactions on targeted clusters 2. On cluster 2, which is Diverse Afternoon Buyers, there are 4190 transaction data. Before the calculation of the association rules with the FP-Growth algorithm, the transaction data has been processed and reduced to prepare the data to be used in the analysis. The process of calculating association rules is done using the Python programming language with the help of the mlxtend library. Data processing to find the association rule is done by comparing the minimum support and the results are shown in [Table 2](#).

Table 2. Support and confidence value comparison

Min support	Min confidence	Number of rules	Biggest lift ratio
0.001	0.4	104	206.9
0.002	0.4	40	112.2
0.005	0.4	8	5.2
0.01	0.4	5	4.8

From [Table 2](#), it is known that the use of different minimum support results in a different number of rules. In this analysis, the researchers used four minimum support values, 0.001, 0.002, 0.005, and 0.01, using a minimum constant confidence of 0.4. The comparison results showed that researchers obtained the largest number of rule and lift ratio at a minimum support of 0.001 with 104 rules and an elevation ratio of 206.9. So, in this study, the minimum support will be used at 0.001 in the analysis of association rules using the FP-Growth algorithm. Using a lower minimum support value allows to identify of more association patterns that may be hidden in the transaction data. It can provide a deeper insight into customer purchasing habits within the "Diverse Afternoon Buyers" cluster and enable the determination of more appropriate and effective

marketing strategies to increase the value of transactions and the profitability of stores. Table 3 shows some important rules that result from the AR-MBA.

Table 3. AR results

No	Rule	Support	Confidence	Lift
1	{Instant fried noodle special} ==> {egg}	0.0360	0.402	1.982
2	{Instant fried noodle special, Instant fried noodle "aceh"} ==> {egg}	0.0026	0.407	2.011
3	{Instant fried noodle on glass} ==> {Instant fried noodle special, Instant fried noodle "aceh", egg}	0.0017	0.412	115.020
4	{soybean sauce} ==> {egg}	0.0012	0.417	2.056
5	{chocolate milk} ==> {chocolate wafer}	0.0012	0.417	96.991
6	{cigarette} ==> {egg}			

It can be seen from Table 3 that eggs have a fairly close relationship with some products such as some kinds of instant fried noodles, soybean sauce, and even cigarettes.

4. DISCUSSION

The results obtained from the data processing of store transactions with cluster analysis and ARMBA can be used as discussion material in developing sales enhancement strategies. The recommendations given in this study focused on promotional strategies that stores could implement, such as product bundling, loyalty points, and discounts. The decision not to recommend changes in the store layout was based on the identification result during the interview process with the store owner. The store owner clearly stated that he was satisfied with the existing store layout and didn't feel there was any significant problem. Therefore, considering direct input from the store owners, this study decided not to propose changes in the store layout. Instead, it is recommended to focus on developing promotional strategies that can improve overall customer performance and loyalty. It is expected to help the store to better results in its business. From the analysis that the researchers have done, then researchers can give recommendations of sales improvement strategies to the store. Some recommendations for the store are as follows,

- Based on four clusters that have been formed using the K-Means algorithm, the researchers targeted cluster 2 as a priority strategy for increasing sales in stores. The selection of the second cluster was based on a high transaction value and a fairly high number of product variations. In addition, it was found that the cluster was actively trading in the timeframe between 13:00 and 19:00, with peak transactions occurring around 14:00 to 15:00. Therefore, the researchers recommended that stores focus sales enhancement strategies to customers in clusters 2 by promoting product bundling at peak times for popular products in the cluster, such as Instant fried noodle special with egg. The discount given may be based on the margin of the bundling product, for example, by giving a discount of 5% of the profit margin. Based on Sheng et al. [27] the use of product bundling and giving discounts to customers can have a positive effect. A bundle price discount can improve consumer positive evaluation of the bundle, increase sales, and attract customers. This is because consumers see the bundles as offers that give more value than buying the products individually.
- The store may provide a point program or a shopping voucher to the customer. Point program can be earned when customers shop with a minimum purchase of Rp. 170.000 and the fold. This figure is obtained from the average purchases made by customers in cluster 4. The point program aims to encourage customers belonging to cluster 2 to become customers entering cluster 4 ("High-Value Customers"). This statue was also performed by Rahmattullah & Yanti [28].
- The store can provide special shelves showing products that have been designated to be bundled in front of shops or near the cashier area. The aim is to raise awareness of product appearance, which can help increase sales and strengthen bundling strategies. Based on Faisal [29] these things simultaneously influence impulse buying decisions so that they can increase sales.

5. CONCLUSION

From the study that has been conducted, it can be concluded that 4 clusters with their specific characteristics can be generated from K-means algorithm. The study focuses on cluster 2 named Diverse Afternoon Buyers because the cluster consists of potential customers who buy varied products with medium total transactions in the afternoon. The transaction in cluster two, then further analysis using AR-MBA, and it can be concluded that egg has an association with several products such as instant fried noodles and cigarettes. The results from both methods were then used to design a sale strategy to increase total transactions, especially in cluster 2. Some of the strategies that are suitable for the store are bundling products, discounts, and vouchers. For clustering analysis, this study only uses attributes that are only provided by the transaction data, so it only can show clusters with limited characteristics. Further study should use more attributes to get more specific customer characteristics.

ACKNOWLEDGMENT

We would like to express our sincere gratitude to ABC store for its invaluable contribution in providing the essential data used in this study. The cooperation and support have played a pivotal role in the success of our endeavors.

REFERENCES

- [1] M. A. Rizaty, "Jumlah Toko Retail di Indonesia Sebanyak 3,98 Juta pada 2022." Accessed: Nov. 06, 2023. [Online]. Available: <https://dataindonesia.id/industri-perdagangan/detail/jumlah-toko-retail-di-indonesia-sebanyak-398-juta-pada-2022>
- [2] F. Achmad, O. Nurdiawan, and Y. A. Wijaya, "Analisa Pola Transaksi Pembelian Konsumen pada Toko Ritel Kesehatan Menggunakan Algoritma FP-Growth," *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 1, pp. 168–175, 2023.
- [3] E. Fitriyani, "Hippindo Ungkap pemicu gerai ritel berguguran: Tempat kegedean, Kumparan." Accessed: Dec. 03, 2023. [Online]. Available: <https://kumparan.com/kumparanbisnis/hippindo-ungkap-pemicu-gerai-ritel-berguguran-tempat-kegedean-1zyVPyvpBIH/3>
- [4] N. H. Tamara, "Perubahan Peta Persaingan Bisnis Retail di Indonesia." [Online]. Available: <https://katadata.co.id/analisisdata/2019/07/03/perubahan-peta-persaingan-bisnis-retail-di-indonesia>
- [5] K. Auliasari and M. Kertaningtyas, "Penerapan Algoritma K-Means untuk Segmentasi Konsumen Menggunakan R.," *Jurnal Teknologi & Manajemen Informatika*, vol. 5, no. 1, 2019.
- [6] Martina, "7 Strategi Pemasaran Pada Bisnis Retail (Eceran)." [Online]. Available: <https://ukirama.com/blogs/7-strategi-pemasaran-pada-bisnis-retail-eceran>
- [7] IBM, *Analytics: The real-world use of big data in retail*. IBM Global Business Services, 2018.
- [8] K. Marwazia Shaliha, A. Angelyna, A. Aulia Nugraha, M. Humam Wahisyam, and T. Kurnia Sandi, "Implementasi K-Means Clustering pada Online Retail berdasarkan Recency, Frequency, dan Monetary (Implementation of K-Means Clustering in Online Retail based on Recency, Frequency, and Monetary)," *Gunung Djati Conference Series*, vol. 3, 2021, [Online]. Available: <https://conferences.uinsgd.ac.id/gdcs>
- [9] A. Sani, "Penerapan Metode K-means Clustering pada Perusahaan," *Jurnal Ilmiah Teknologi Informasi*, vol. 1, no. 7, 2018.
- [10] I. Kamila, U. Khairunnisa, and Mustakim, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, pp. 119–125, 2019.
- [11] M. Sarkar, A. R. Puja, and F. R. Chowdhury, "Optimizing Marketing Strategies with RFM Method and K-Means Clustering-Based AI Customer Segmentation Analysis," *Journal of Business and Management Studies*, vol. 6, no. 2, pp. 54–60, 2024.
- [12] F. Marisa, S. S. S. Ahmad, Z. I. M. Yusof, F. Hunaini, and T. M. A. Aziz, "Segmentation Model of Customer Lifetime Value in Small and Medium Enterprise (SMEs) using K-Means Clustering and LRFM Model.," *International Journal of Integrated Engineering*, vol. 11, no. 3, 2019.

- [13] Y. Li, X. Chu, D. Tian, J. Feng, and W. Mu, "Customer segmentation using K-means clustering and the adaptive particle swarm optimization algorithm," *Appl Soft Comput*, vol. 113, 2021.
- [14] R. Siagian, P. Sirait, and A. Halima, "E-Commerce Customer Segmentation Using K-Means Algorithm and Length, Recency, Frequency, Monetary Model," *Journal of Informatics and Telecommunication Engineering*, vol. 5, no. 1, 2021.
- [15] M. Kaur and S. Kang, "Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining," in *Procedia Computer Science*, 2016, pp. 78–85. doi: 10.1016/j.procs.2016.05.180.
- [16] Y. P. Bunda, "Algoritma Fp-Growth Untuk Menganalisa Pola Pembelian Oleh-Oleh (Studi Kasus Di Pusat Oleh-Oleh Ummi Aufa Hakim)." *Riau Journal of Computer Science*, vol. 6, no. 1, 2020.
- [17] A. Setiawan, D. Indra, and G. Anugrah, "Penentuan Pola Pembelian Konsumen pada Indomaret GKB Gresik dengan Metode FP-Growth," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 2, no. 2, 2019.
- [18] Erwin, "Analisis Market Basket Dengan Algoritma Apriori dan FP-Growth," *Jurnal Generic*, vol. 2, no. 2, pp. 472–478, 2009, Accessed: Dec. 10, 2023. [Online]. Available: <http://generic.ilkom.unsri.ac.id/index.php/generic/article/view/15>
- [19] D. Dwiputra, A. M. Widodo, H. Akbar, and G. Firmansyah, "Evaluating the Performance of Association Rules in Apriori and FP-Growth Algorithms: Market Basket Analysis to Discover Rules of Item Combinations," *Journal of World Science*, vol. 2, no. 8, 2023.
- [20] A. H. Wibowo, H. Windyatri, A. Marulan, D. Harits, and Satria, "A Pattern Analysis on Goods Purchase Relationship for Minimarket's Customers by Using Association Rule - Market Basket Analysis Method (AR-MBA)," *Journal of Information Technology Ampera*, vol. 4, no. 3., 2023.
- [21] I. D. Putri, Yuhandri, and R. Hardianto, "Application of the FP-Growth Algorithm in Consumer Purchasing Pattern Analysis," *Journal of Computer Science and Information Technology*, vol. 10, no. 2, pp. 44–49, 2024.
- [22] L. Hao, T. Wang, and C. Guo, "Research on parallel association rule mining of big data based on an improved K-means clustering algorithm," *Int J Auton Adapt Commun Syst*, vol. 16, no. 3, pp. 233–247, 2023.
- [23] L. Kusak, F. B. Unel, A. Alptekin, M. O. Celik, and M. Yakar, "Apriori association rule and K-means clustering algorithms for interpretation of pre-event landslide areas and landslide inventory mapping," *Open Geosciences*, vol. 13, no. 1, pp. 1226–1244, 2021.
- [24] S. M. Dol and P. M. Jawandhiya, "Classification Technique and its Combination with Clustering and Association Rule Mining in Educational Data Mining — A survey," *Eng Appl Artif Intell*, vol. 122, 2023.
- [25] M. M. Rahman, M. S. Islam, R. R. Richi, and A. Chakraborty, "Course Recommendation System for Students Using K-Means and Association Rule Mining," in *2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Ankara, Turkey: IEEE, 2022, pp. 641–646.
- [26] S. Fitri Octavia, I. Permana, and S. Monalisa, "JURNAL MEDIA INFORMATIKA BUDIDARMA Penerapan Algoritma Association Rules Dalam Penentuan Pola Pembelian Berdasarkan Hasil Clustering," 2023, doi: 10.30865/mib.v7i3.6129.
- [27] S. Sheng, A. M. Parker, and K. Nakamoto, "The effects of price discount and product complementarity on consumer evaluations of bundle components," *Journal of Marketing Theory and Practice*, vol. 15, no. 1, pp. 53–64, 2007.
- [28] R. Rahmattullah and R. Yanti, "Peningkatan Strategi Penjualan dengan metode Association Rule Pada Toko Ritel ABC," 2022.
- [29] I. Faisal, "Pengaruh Price Discount, Bonus Pack, Dan In-Store Display Terhadap Keputusan Impulse Buying Pada Giant Ekstra Banjar.," *Jurnal Sains Manajemen Dan Kewirausahaan*, vol. 4, no. 2, pp. 266–274, 2018.