Journal of Applied Data Sciences
Vol. 6, No. 1, January 2025, pp. 86-100

Designing Planograms for Retail Shelves: A Visual Merchandising Approach Using Apriori Algorithm and K-Means Clustering of Customer Preferences

Jundi Nourfateha Elquthb[1], Agus Mansur[2,*]

[1, 2]*Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia, Yogyakarta 55584, Indonesia*

(Received: August 14, 2024; Revised: September 5, 2024; Accepted: October 24, 2024; Available online: December 27, 2024)

Abstract

The development of retail store businesses in Indonesia is widespread across various regions, positively impacting the increase in shopping activities among the community. Alongside the intense competition among retailers, fundamental issues are found regarding the shopping experience of customers who feel uncomfortable and dissatisfied with retail services. Customers are an essential aspect that needs attention, and their desires must be fulfilled to maintain the existence of a retail store. This study attempts to implement the concept of visual merchandising to enhance service quality through the planogram method aimed at improving the visual arrangement of sales shelfs. Regarding the layout of retail stores, shelf arrangement plays a significant role visually in influencing and attracting customer attention while shopping. In this study, two data mining techniques are used. The first method is association rule mining using the Apriori algorithm, which reveals the association rules formed between two or more product items, utilizing a total of 6,325 customer transaction records. The results indicate 12 rules formed based on product categories and 17 rules based on product sub-categories. The second technique is k-means clustering, which is used to identify differences in customer preferences in retail stores based on several variables regarding customers using 104 data customers. In practice, both the apriori and k-means algorithms face challenges, especially when handling large and complex data. Differences in preferences were found among the clusters, particularly regarding product arrangement on the sales shelf, such as grouping products by brand and price. A two-dimensional planogram design was developed to integrate the results from the previous stages. This design considers the availability and specifications of the shelves in the retail store. A planogram adjusted to customer purchasing patterns and preferences is expected to provide good service and positively impact store operations and stock management.

Keywords: Association Rule, Customer Preferences, K-Means Clustering, Planogram, Retail Shelf, Visual Merchandising

1. Introduction

The sustained impact of various sectors, such as the economy and society, has driven significant economic growth in Indonesia. These two sectors are closely intertwined with changes in human lifestyle patterns, which can boost buying and selling activities in a market. One form of economic advancement can be seen in the proliferation of retail stores spread across Indonesia. The presence of these retail stores provides significant benefits, especially for the community, by making it easier for customers to meet their needs. This retail store development has a positive impact, as household consumption dramatically contributes to Indonesia's gross domestic product (GDP) growth.

However, retail managers still face crucial problems in providing services. This situation occurs due to the uneven implementation of management standards by retail managers. One factor is that these retail stores are still considered traditional due to limited information, physical environmental conditions, and capital resources [1]. The retail management process becomes less optimal, particularly in providing services that satisfy customers and draw attention to the products being sold. Mistakes in product placement on shelves are also a problem that can lead to low customer satisfaction in a retail store. Consequently, customers may have difficulty finding their desired products and complain about product placement issues and inaccurate information [2]. There is a relationship between customer shopping

***Corresponding author: Agus Mansur (agusmansur@uii.ac.id)**
DOI: https://doi.org/10.47738/jads.v6i1.375

experiences and the diverse conditions of retail stores. If a retail store has clear visibility and shopping flow, it will affect its attractiveness to customers. The more appealing the visual display and impression the store management accommodates, the higher the likelihood of increasing sales and competing effectively with rivals. If customers negatively perceive the store's layout and visual aspects, it can decrease their interest in shopping [3]. In addition to managing the retail visual display, understanding customer purchasing patterns is crucial to retail success. Information about purchasing patterns and product preferences can serve as a basis for implementing better sales strategies [4]. Furthermore, it is essential to identify customer preferences that reflect the shopping experience they desire. Addressing both aspects makes it possible to resolve customer complaints through marketing strategies that offer the best deals and services [5].

The approach and application of data mining techniques can be utilized to accommodate valuable information related to improving customer service. The association rule method is used to identify customer purchase patterns based on historical sales data from retail stores. The results reveal a strong relationship between two product items likely to be purchased together in a single transaction [6]. Products involved in association rules indicate a high frequency of simultaneous purchases and should be further examined by retail store managers. Identifying customer preferences uses the K-Means clustering method to understand customer characteristics. With clustering, several smaller groups can be identified, revealing differences in customer desires or behaviours between the formed clusters related to their shopping experience in the retail store.

Customer shopping experience is crucial and is influenced by the visual aspects of the store. The better a retail store's visual management, the more it will enhance customer attraction during shopping and provide satisfaction with the service [7]. This research aims to accommodate all findings from data mining techniques and develop them through visual merchandising in retail stores, explicitly focusing on shelf display. This development uses the planogram method, which focuses on arranging product displays on shelves to attract customer attention. The application of a planogram for shelf displays needs to consider several critical visual aspects related to the products, such as quantity, size, volume, and other specifications [8]. Implementing this planogram is hoped to serve as a basis for retail store managers to adjust store conditions by applying visual merchandising. Both data mining methods used in this study aim to explore further and uncover patterns formed from key elements in retail business operations. There is a close relationship between product sales and customer behavior, which can be combined through the chosen data processing methods. The patterns or knowledge discovered are highly valuable for business development in retail, particularly in terms of marketing that relies on visual aspects. This concept can be used to increase product attractiveness to customers, making products more accessible to recognize and encouraging higher purchase intentions.

2. Literature Review

2.1. Retail Store

Retail is a business sector that provides and sells goods or services in individual units to customers. The products or services sold are intended for personal use by the end customer, not for further business purposes [9]. Retail development has introduced various types of retail stores, leading to intense competition among retail managers. Effective retail management is crucial to maintaining the business's presence by ensuring smooth product distribution and quality service. The success of a retail store is greatly influenced by its customers, who are the primary sales targets. Additionally, it is important to understand customer needs and behaviors to ensure their comfort and positively impact the store's existence through increased customer loyalty [10]. The important aspect of retail service is the store layout, which plays a crucial role in the effectiveness of the shopping process. Providing an appealing visual impression, organizing sales products, and ensuring smooth and comfortable customer traffic is essential [11]. Customers will explore aisles and navigate the store to find their desired products. Therefore, retail managers should provide clear information to assist customers, as it is common for customers to make impulse purchases after browsing the store. A well-designed store layout can create a positive shopping experience and increase product sales. The store layout is adjusted to the manager's needs and sales strategy, while considering cleanliness, security, and strategic product placement [12].

ISSN 2723-6471

2.2. Visual Merchandising

Customers will feel satisfied and comfortable when they have a positive shopping experience at a retail store. A good shopping experience can stimulate customers to shop again, facilitate finding products, and increase their shopping time at the store [13], [14]. Providing services can be enhanced by integrating visual aspects and store conditions into sales strategy designs, representing non-verbal communication to customers. One such strategy is visual merchandising, which combines applied science and art to create an appealing presentation of products. Visual merchandising can influence customers to purchase products by considering several visual attributes such as store condition, product value, and product quality [15]. Important aspects include the store's exterior, sales displays, atmosphere, and promotional advertising. These attributes and aspects are used in visual merchandising to provide information on product specifications, location, customer circulation flow, and responsive service [16].

2.3. Planogram

A planogram is a design concept for layout and display that considers the conditions of a retail shelf and the availability of products in a retail store. It is often referred to as space management because its purpose is to manage the arrangement and allocation of products comprehensively to attract customer attention and positively impact sales [17]. The development of a planogram can be carried out using previous evaluations, such as sales data and inventory management in a retail setting while considering various aspects like the ideal number of items per shelf row, product and shelf specifications, and the art of combining products with others [8]. In the retail business, planogram design is executed as effectively as possible, maintaining neatness by grouping products based on size, weight, type, and color on specific shelf displays [3].

2.4. Apriori Algorithm

The Apriori algorithm is one of the algorithms used in association rules, designed to discover relationships or association rules within large datasets. The application of association rules can identify an event along with its consequences in an IF-THEN logical statement. In association rules, there are three primary parameters: support, which describes the frequency of an item in the dataset; confidence, which describes the certainty of a rule or relationship between items that are formed; and lift ratio, which determines the validity of a rule, where the ratio must be greater than 1 [6]. Apriori can be used for the process of identifying association rules using two techniques: pruning and data merging. The pruning technique is used to reduce the search space based on the support and confidence values in the processed data. Meanwhile, the merging technique is done by combining the rules formed from the pruning results to find valid association rules [18]. According to [19], Apriori has the following main stages: determining the support value, calculating candidate itemsets, recalculating candidate itemsets, and repeating the process to ensure the final result (i.e., no additional itemsets remain). The application of association rules or the Apriori algorithm is typically used for market basket analysis and banking transactions, which can be leveraged as a strategy management tool for improvement [20].

2.5. K-Means Clustering

Clustering is one of the critical functions in data mining, aimed at grouping data into specific clusters. The data being grouped is taken from an extensive database and then processed through the clustering stage to identify similarities in characteristics and proximity between objects [21]. In clustering, there are two types of data processing: hierarchical and non-hierarchical. One of the commonly used algorithms in non-hierarchical clustering is the K-Means algorithm, which allows for determining the number of clusters through trial and error or according to the research objectives [22]. However, to determine the ideal number of clusters, cluster data testing is required using the elbow and silhouette visualization methods. In the elbow method, the number of clusters can be identified by observing the percentage comparison results, which will form a sharp angle at a certain point (resembling an elbow) [23].

ISSN 2723-6471

3. Methodology

Several key stages are involved in addressing the problem and achieving the objectives of this research. These stages are interconnected in designing the outcome, such as a planogram for shelf sales displays in retail stores. The flow and mechanism of this research are illustrated in figure 1 below.

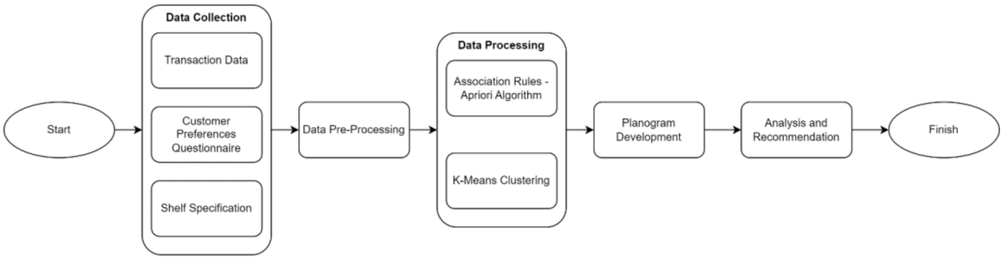

Figure 1. Research Stages

Several methods are employed, including the Apriori algorithm, K-means clustering, and planogram. Each method serves a specific function in formulating and solving the problem of visual merchandising for retail store shelves. The use of the Apriori algorithm in the application of association rule methods is due to its superiority in handling large datasets, allowing it to optimally search for item combinations in high-volume sales transactions [24]. Then, K-means clustering is used to determine customer clusters based on predefined characteristics. This algorithm can obtain results through an iterative approach and divides objects into various clusters by adjusting the Euclidean distance to the centroid centers of the involved objects [25]. Meanwhile, the planogram is used to accommodate all findings from the two data mining techniques, with improvements made from a visual aspect perspective. This method is generated in the form of identifying the design of sales shelves, adjusted to the purchase patterns of items and customer preferences.

3.1. Data Collection

Three methods require data collection: association rules, clustering, and planogram. In the association rules method (Apriori algorithm), the data used are sales transaction records from a retail store located in Yogyakarta. Over one month, data were collected from 6.325 customer transactions, which were grouped based on transaction numbers. For the clustering method, data related to customer characteristics were obtained through questionnaires distributed to retail store customers. The customer data encompasses several aspects that support this research, including demographics, visual merchandising evaluation, customer satisfaction and loyalty, and shelf design preferences, as shown in table 1 on the determination of research aspects and variables below.

Table 1. Aspects and Variables of Questionnaire Model

Aspect	Variable	Code
Demography	Gender	X1
	Age	X2
	Monthly Income	X3
	Shopping Expenses	X4
Visual Merchandising [26]	Store Exterior	X5
	Store Interior	X6
	Interior Point of Purchase Display	X7
	Store Layout and Shelf	X8
Customer [27]	Satisfaction	X9
	Loyalty	X10

Aspect	Variable	Code
Planogram Design [3]	Product Brand	X10
	Colour	X11
	Size	X12
	Price	X13

Data was collected from 104 respondents who are customers of the retail store. The data used in the clustering process adequately represents the population of customers shopping at the retail store, which falls into the category of small to medium-sized retail stores. The shelf planogram development method utilized findings from association rules and clustering methods. The results from association rules help identify which products are potential candidates for inclusion in the planogram design strategy. The clustering method provides information about customer interests and shopping characteristics, specifically related to visual merchandising aspects and preferences for shelf sales display design. Additionally, data collection included specifications for shelf displays, shelf arrangement conditions, and available shelf layouts in the retail store. Various shelves with different specifications, including height, width, and depth, were considered. The planogram integrates these differences to create a shelf display design tailored to customer preferences.

3.2. Data Pre-Processing

The collected data then proceeds to the next stage, data pre-processing, including cleaning up double entries, noise, and invalid data. In association rules, duplicate data refers to items in a transaction that share the same category, where one item is used as a representative. Additionally, noise and invalid data are removed, which includes data that has no correlation with the study, such as input errors made by respondents. In the association rules method, the sales transaction data used are transactions that include two or more product items. The data is then transformed by changing the naming conventions based on product categories and subcategories. This grouping process follows the designation of 15 product categories and 65 sub-categories. After data transformation, 3.242 transactions were selected based on categories, and 3.655 valid transactions were selected based on subcategories. There was a decrease in the number of transactions due to this pre-processing and differences in the product grouping rules. Additionally, customer questionnaire data undergo a transformation process using the Likert scale or weighting for specific questions. This transformation aims to standardize the collected data into a simpler and more calculable form for the K-means clustering method.

3.3. Retail Transaction Data

Processing association rules involves using transaction data selected based on their categories and subcategories. At this stage, the Apriori algorithm is applied to generate association rules that may emerge from the transaction data. Before processing the transaction data, it is necessary to determine the minimum values for the association rules parameters, namely support and confidence. Researchers have the flexibility to set these parameter values according to their needs or the characteristics of the data [28]. Therefore, experiments with different parameter scenarios can be conducted to determine the appropriate parameter values, taking into account the number of output rules that can be generated, as shown in table 2 and table 3 below.

Table 2. Parameter Determination for Product Category

Trial	Support	Confidence	Output Rules
1	30%	50%	0
2	15%	30%	9
3	5%	25%	28

ISSN 2723-6471

Table 3. Parameter Determination for Product Sub-category

Trial	Support	Confidence	Output Rules
1	7.5%	20%	0
2	4%	20%	4
3	3%	20%	19

The tables above are the result of trial and error to determine the minimum parameter values (support and confidence) used, based on the number of outputs generated and having a lift ratio value above 1. The results indicate that the smaller the minimum support value, the more rules can be generated. The selection of support and confidence thresholds in this study takes into account the total amount of data obtained and processed. The data used consists of thousands of transactions, containing a wide range of product categories. The more products that appear in a transaction, the smaller the support value used, which is proportional to the product frequency ratio. Similarly, the confidence value reflects the certainty of the rules formed, and this determination is well-suited for retail store transaction data due to the large volume of data involved.

3.4. Customers Preferences Questionnaire

A total 104 retail store customers answered all the questions distributed through a digital questionnaire related to the specified aspects. One specific variable in the questions was related to the customers preferred eye-catching colors. Customers have particular tendencies and interests when identifying products based on color. As shown in figure 2 below, the customer color preference response is presented.

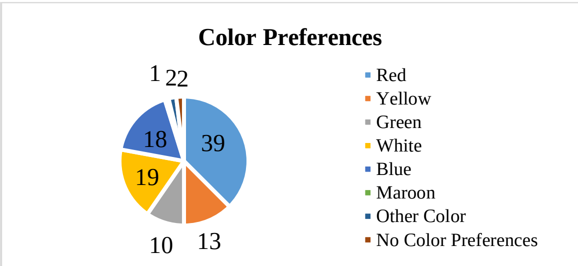

Figure 2. Color Preferences

The results from the above image reveal that the most frequently chosen colors are red (39 customers), white (19 customers), and blue (18 customers). Based on these preferences, these three colors will be used as the foundation for designing a planogram that aligns with the majority preferences. Proper use of color in product placement on shelf displays can enhance visual appeal and make it easier for customers to find the products.

The multicollinearity test aims to identify any high correlation between variables, as a significant correlation can disrupt the stability of the final clustering model results. Multicollinearity among independent variables has a detrimental impact on the interpretation process of clustering results because it can create an imbalance in weighting, which in turn affects the linear regression estimator, making it biased and inaccurate according to predictions [29]. In the context of customer-related variables, a high correlation between independent variables can lead to unclear cause-and-effect relationships. For example, when a variable has a very high correlation with two other variables, it becomes difficult to determine the primary reason why a customer behaves or possesses unique characteristics. As a result, it becomes challenging to develop interpretations to understand customer characteristics and to identify the reasons or preferences behind the customer's behavior [30]. Figure 3 below presents the results of multicollinearity testing for all variables using RStudio software.

ISSN 2723-6471


```
> print(vif_values)
         Age       Income        Spent     Exterior     Interior      Display       Layout
    1.161506     1.448189     1.229284     3.808745     2.699231     2.735498     2.423061
Satisfaction     Decision        Brand        Color         Size        Price
    1.705883     1.326200     1.155838     1.114958     1.196385     1.132260
```

Figure 3. Multicollinearity Test

According to the rule that a variable is free from multicollinearity if it has a Variance Inflation Factor (VIF) value of less than 10, the test results indicate no significant multicollinearity in this clustering model. This result is shown by the fact that the VIF values for all variables are below 10, allowing for the continuation of the next stage of clustering. Data processing to determine the optimal number of clusters is shown in figure 4 and figure 5 below for each method.

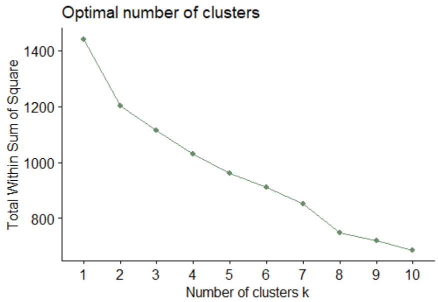

Figure 4. Elbow Method

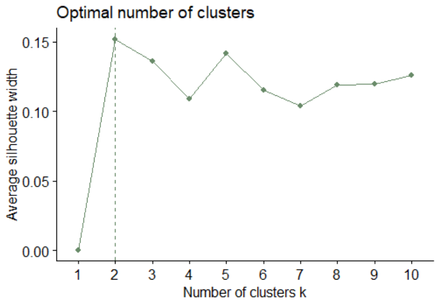

Figure 5. Silhouette Method

Data cleared of assumption tests undergoes a transformation known as data standardization. This transformation process aims to normalize the range of numerical data to be more uniform, facilitating the clustering of data processing. This study employed a visual approach using silhouette and elbow methods to determine the ideal number of clusters to form. The elbow method identifies the ideal number of clusters based on the first significant bend (similar to an elbow) at 2. In contrast, the silhouette method is determined by the vertical line that appears at 2. The selection of two clusters formed aims to test the representativeness of the sample population, which is assumed to be divided into different groups. This approach is facilitates the analysis of differences in the average value or characteristics of the customers within each cluster. Subsequently, statistical approaches such as T-tests or F-tests may be used, depending on the applicable case study. However, in this research, the analysis is limited to generating customer groups, which are then interpreted using the average values of each variable for profiling purposes.

3.5. Retail Shelf Specification

This study utilizes supporting data in the form of shelf sales display specifications as a consideration in the planogram design development process. The retail store contains various types of shelves, each with its dimensions, as detailed in table 4 below regarding the specifications of shelf sizes. It was found that the products on the shelf sales display are arranged in such a way by the retail managers to attract customer attention.

Table 4. Retail Shelf Specification

Types	Size (cm)			
	Height	Width	Row Height	Depth
Cashier	90	55	20	20
	90	160	20	20
Small	137,5	75	15	30
	137,5	75	17	30
Medium	135	90	21	30
	135	90	30	30

Types	Size (cm)			
	Height	Width	Row Height	Depth
Large	180	87,5	25	30
	180	87,5	40	30
Refrigrator	180	125	28	50

The variety of shelf types available in retail stores provides insight and an advantage for this research. It can be used to address and resolve issues related to product size preferences in a shelf sales display and as a platform for more optimal product management. The diversity of types and sizes of shelves presents its own challenges for retail store managers or merchandisers managing product display shelves. Therefore, improvements are needed, including a combination of structure and shelf size customization to accommodate or address issues related to visual merchandising concepts.

4. Results and Discussion

4.1. Association Rules – Apriori Algorithm Results

The results were obtained by determining or experimenting with the minimum values of the support and confidence parameters. The study indicates that the higher the support and confidence values, the more likely these rules will apply in a transaction. According to the established criteria, a rule is considered valid if it has a lift ratio greater than one and includes at least two different product items. During the data processing stage for product categories, 28 rules were generated, and after sorting, only 12 valid rules were identified. Consequently, 19 rules were generated for product subcategories, and 17 valid rules were selected, as shown in table 5 and table 6.

Table 5. Categories Rules

No.	rules	support	confidence	coverage	lift	count
1	{Groceries} => {Breakfast}	0.074	0.432	0.171	1.198	239
2	{Breakfast} => {Groceries}	0.074	0.204	0.361	1.198	239
3	{Breakfast, Confectionery} => {Food}	0.056	0.679	0.082	1.143	180

Table 6. Sub-categories Rules

No.	rules	support	confidence	coverage	lift	count
1	{Chips} => {Cake n Bread}	0.033	0.253	0.131	1.862	121
2	{Cake n Bread} => {Chips}	0.033	0.243	0.136	1.862	121
3	{Sachet Drink} => {Instant Noodle}	0.035	0.231	0.152	1.63	128

Table 5 and table 6 above present the results of sales transaction data processing using the apriori algorithm approach. In this study, 3 rules were showed with the highest lift ratio values in both category and sub-category scenarios from the dozens of rules formed. Based on the results, rules with high confidence and lift ratio values indicate a stronger andmore valid likelihood of a relationship between two or more item combinations. Figure 6 and figure 7 illustrate the results of visualizing the relationships between items or product categories that emerged after data processing.

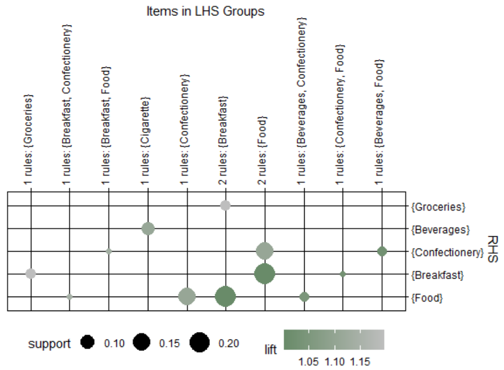

Figure 6. Rules Results Based on Product Category

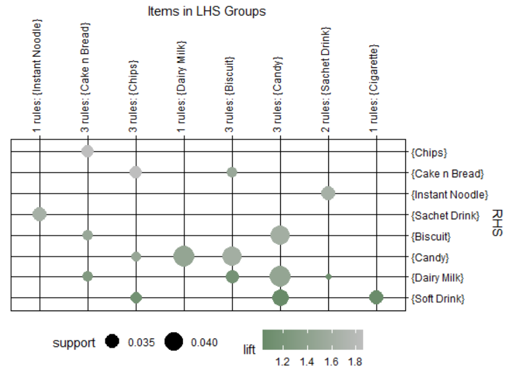

Figure 7. Rules Results Based on Product Sub-category

The items on the left-hand side (LHS) represent the primary items (antecedents) that can trigger the purchase of items on the right-hand side (RHS), which are secondary (consequents). This association rule is visualized in the form of a matrix group where the larger the circle, the higher the support value, and the darker the green color of the circle, the higher the confidence value of the rule. Most results from both scenarios (category and sub-category) of transaction data processing show that the items involved in the association rules belong to categories of consumable goods. This finding means that these products are frequently purchased by customers and categorized as Fast Moving Consumer Goods (FMCG), which have a short shelf life due to high consumer demand. Therefore, retail managers must pay attention to the inventory and placement of certain products to ensure customers easily find them, especially for daily consumable products (food and beverages) are those with high confidence values.

Retail store managers can implement several suggestions regarding layout, such as providing sufficient space for customer movement. Adequate spacing or distance between shelf displays will create a comfortable shopping experience and provide insights into customer movement patterns during shopping [31]. Grouping products by category enhances customer comfort by ensuring that products are not contaminated by one another [32]. Specifically, this study uses a combined approach to incorporate key findings such as store conditions, room size, shelf spacing, and product positioning into a planogram method. A planogram can be a strategic tool to increase customer satisfaction, attract attention, and stimulate purchases.

4.2. K-Means Clustering Results

In the previous stage, it was determined that based on the elbow and silhouette methods, the optimal number of clusters formed was two using the K-Means method. With the formation of these two clusters, further analysis can be conducted to explore whether there are differences in characteristics among these customers. The determination of the number of clusters formed using silhouette and elbow methods aims to ensure that the clusters are optimal in terms of both quantity and membership. Figure 8 illustrates the visualization of customer distribution based on the clustering results.

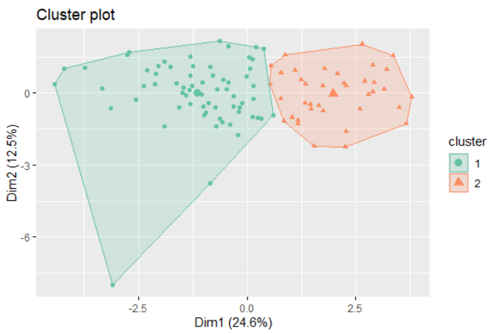

Figure 8. K-Means Cluster Plot

The optimal number of clusters is determined through the silhouette score and elbow method, which indicate the quality of the formed clusters [33]. The results show that Cluster 1 has 66 customers, while Cluster 2 has 38 customers. In Fig.

8, the customer distribution visualization based on Cluster 1 is depicted in green, and Cluster 2 is depicted in orange color. Table 7 displays the profiling results of the established clusters according to the data transformation rules.

Table 7. Cluster Profiling

Variable	Cluster Average		Cluster Profile	
	1	**2**	**1**	**2**
X1	1.53 (2)	1.47 (1)	Women	Men
X2	23.56 (24)	25.76 (26)	24	26
X3	2.43 (2)	3.00 (3)	500.000 – 1.000.000	1.000.000 – 2.000.000
X4	2.90 (3)	3.05 (3)	50.000 – 100.000	50.000 – 100.000
X5	13.84 (14)	17.50 (18)	Fair	Good
X6	10.47 (10)	13.26 (13)	Fair	Good
X7	10.16 (10)	12.76 (13)	Fair	Good
X8	10.00 (10)	13.34 (13)	Fair	Good
X9	3.75 (4)	4.50 (5)	Satisfied	Very Satisfied
X10	1.98 (2)	2.00 (2)	Return	Return
X11	1.22 (1)	1.55 (2)	Grouped	Blended
X12	1.25 (1)	1.26 (1)	Grouped	Grouped
X13	1.18 (1)	1.02 (1)	Sorted	Sorted
X14	1.72 (1)	1.42 (1)	Sorted	Blended

The result shown in table 7 that the characteristics were determined based on the average values of most cluster members (customers), followed by rounding and adjustments to the research model transformation. There are significant differences between the two customer clusters across several variables. Cluster 1 is dominated by male customers with lower spending per shopping trip, while female customers with higher spending and higher satisfaction levels dominate Cluster 2. For the design of sales shelf planograms, the differences between these clusters regarding product grouping by brand and price indicate that shelf design preferences vary. These preference differences should be considered in planogram development to meet the needs of each cluster. In terms of customer satisfaction, it has a positive impact on the level of customer loyalty in retail stores. In this study, it was found that the majority of customers are loyal and will return to the retail store for future shopping. Customer satisfaction and loyalty are significantly influenced by the perceived visual atmosphere of the retail store. Therefore, it is crucial to organize and manage the retail store's visual layout in aspects such as product facing, product positioning (vertical and horizontal), category arrangement, and product adjacencies. By integrating these aspects, a visual value can be created that enhances the effectiveness of the store layout and increases customer attractiveness during shopping [34].

Based on gender, clusters of female customers should be targeted with sales strategies that appeal to women and offer visually appealing product arrangements. Women customers make impulsive decisions when purchasing certain products, influenced by the condition or visual appeal [35]. Equally important is product pricing, which relates to the customers' financial considerations. Customers will select products based on their shopping budget, evaluating the price quality concerning the cost [36]. A different approach can be applied to clusters of men customers, who tend to spend less time shopping than women. Men customers are often quick to compare products and make purchasing decisions. Therefore, a suitable strategy would be to place products targeted at men in strategic locations within the retail store and to provide detailed product information [37]. Additional insights related to product color indicate that there is a tendency for color choices to vary based on gender. Preferred colors among customers, both men and women, are blue and red, which are recognized for their universal appeal and easy identification [38].

Based on the cluster characteristic analysis, significant differences were found in several aspects, such as demographics, visual merchandising evaluation, loyalty, satisfaction, and planogram design preferences for shelves. This study highlights how product arrangement on sales shelves through planograms can enhance customer satisfaction by creating visual appeal and making it easier to find products. There is a striking difference between the two clusters regarding design preferences and product arrangement on shelves. Female customers prefer products grouped by brand and organized from the cheapest to the most expensive. In contrast, male customers prefer products not grouped by brand (blended) and do not consider price order. These findings indicate varying preferences among the clusters. Customers will feel satisfied if they can quickly find the products they want. Retail store managers should leverage visual merchandising concepts to enhance this experience, focusing on interior design, visuals, and shelf layout. Effective product arrangement considers factors such as brand, size, color, and price grouping and provides markers that reflect customer preferences. This strategy should be tailored to the purchasing patterns previously identified through the Apriori algorithm to maximize profits and reach a broader customer base.

4.3. Planogram Design Results

The development of shelf design is conducted by considering the research findings using association rules and clustering methods. The product transaction combinations used are those from the food and beverage category because they have a high confidence value. Related products are plotted into the sale shelf display with several adjustments. Customer preference clustering results are also used to determine the product arrangement on the shelf. Customers in the two clusters have different product arrangement preferences, particularly regarding brand positioning and product ordering based on price.

Figure 9 below shows the results from processing and developing a 2D planogram design adjusted to customer preferences. The products displayed represent the items that appear in the association rules, with the condition that they have a high confidence value. Various adjustments are needed regarding the shelf size and the spacing between rows to address the issue of product arrangement based on size. Therefore, different types of products are combined on a single shelf based on their size and weight. The design of this planogram results from implementing visual merchandising concepts, which are essential elements in the product sales strategy on shelves. Retail store managers or merchandisers are responsible for providing clear information on the shelves. The visual design aspect serves as a communication tool between the product provider and the customer, encompassing elements such as text information (price, promotions, bundling), brand label information, and color selection that is easily recognizable to the customer’s eye [39].

The arrangement of products using planograms serves various purposes to meet customer preferences and optimize product allocation on shelves. There is a relationship between brand and price, as customers usually have standards determined by the perceived quality of a brand's product. Product placement on shelves can focus on specific brands with high potential or boost the sales of a specific product [36]. Additionally, ordering price labels on a row of shelves (within the same product category) provides a means for customers to compare products. Comparing products based on price is one way of making purchase decisions, commonly called price awareness about the product's quality [40]. It is crucial to provide price information to communicate with customers, which can influence the arrangement or placement of products based on their prices, as it can serve as an efficient profit analysis tool. Below, figure 9 represents the planogram resulting from the data processing of product item combinations and customer preference clustering.

Figure 9. Shelf Planogram

The combination of shelf spacing applied in this planogram design aims to maximize the allocation or plotting of products. The planogram design and product arrangement must consider the standard shelf specifications based on height, width, depth, and row spacing [41]. A single shelf can hold many products of various sizes and weights. Small and lightweight products are placed on the upper shelves to be easily reached by customers, while larger products are allocated to the lower shelves to be closer to the floor support of the shelf. Small products can be combined on the same shelf row, or a dedicated space can be provided for potential products. This approach is essential to facilitate product identification, improve customer reach, and enhance the neatness and aesthetics of the shelf display [42], [43].

Easily recognizable colors can capture customer attention and potentially boost sales. Therefore, these specifically colored products are placed on the same shelf. Organizing products by color is the most effective communication tool to ensure that the products and packaging are easily visible to customers [44]. However, from the entire application of visual merchandising concepts mentioned above, it would be beneficial if future research could incorporate a process of verification or proof of results. By implementing a mixed strategy or combining product arrangement scenarios, a comparison between certain merchandising scenarios is needed to observe the differences in their effects. This aims to demonstrate how effective the merchandising strategy is in relation to the purchasing rate or the profit gained.

5. Conclusion

The transaction sales data holds valuable insights when processed using the Apriori algorithm, revealing customer purchasing patterns in retail stores. The data analysis results show that numerous association rules have been discovered (12 rules in the category and 17 rules in the sub-category) involving products that are typically daily consumption items frequently purchased by customers. These products require a particular location to ensure their quality is maintained and customers can easily find them. This can be maximized by adjusting and reorganizing the store layout to be managed based on product categories. The findings from the K-Means clustering indicate differences in customer characteristics within each cluster, particularly in aspects related to visual preferences and planogram design, which involve the arrangement of products based on brand, size, color, and price. Organizing products on retail store shelves is crucial for optimizing shelf management. Creating a planogram while considering the visual merchandising concept aims to provide adequate customer service or sales communication. The implementation of planogram design and product arrangement must consider the availability of space and containers in the form of shelves and their detailed

 ISSN 2723-6471

specifications. These adjustments are expected to meet the diverse preferences of customers, ultimately improving the retail store's service quality.

6. Declarations

6.1. Author Contributions

Conceptualization: J.N.E. dan A.M.; Methodology: A.M.; Software: J.N.E.; Validation: J.N.E. dan A.M.; Formal Analysis: J.N.E. dan A.M.; Investigation: J.N.E.; Resources: A.M.; Data Curation: A.M.; Writing Original Draft Preparation: J.N.E. dan A.M.; Writing Review and Editing: A.M. dan J.N.E.; Visualization: J.N.E.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received financial support for the research and publication of this article from the Department of Industrial Engineering, Faculty of Industrial Technology, Universitas Islam Indonesia, Yogyakarta, Indonesia.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] T. Reardon and R. Hopkins, “The supermarket revolution in developing countries: Policies to address emerging tensions among supermarkets, suppliers and traditional retailers,” *Eur. J. Dev. Res.*, vol. 18, no. 4, pp. 522–545, 2006, doi: 10.1080/09578810601070613.

[2] C. H. Lee, X. Zhao, and Y. C. Lee, “Service quality driven approach for innovative retail service system design and evaluation: A case study,” *Comput. Ind. Eng.*, vol. 135, no. February 2018, pp. 275–285, 2019, doi: 10.1016/j.cie.2019.06.001.

[3] K. Czerniachowska, *A genetic algorithm for the retail shelf space allocation problem with virtual segments*, vol. 59, no. 1. Springer India, 2022. doi: 10.1007/s12597-021-00551-3.

[4] E. Bilgic, O. Cakir, M. Kantardzic, Y. Duan, and G. Cao, “Retail analytics: store segmentation using Rule-Based Purchasing behavior analysis,” *Int. Rev. Retail. Distrib. Consum. Res.*, vol. 31, no. 4, pp. 457–480, 2021, doi: 10.1080/09593969.2021.1915847.

[5] J. R. Hanaysha, “An examination of the factors affecting consumer’s purchase decision in the Malaysian retail market,” *PSU Res. Rev.*, vol. 2, no. 1, pp. 7–23, 2018, doi: 10.1108/PRR-08-2017-0034.

[6] W. P. Nurmayanti *et al.*, “Market Basket Analysis with Apriori Algorithm and Frequent Pattern Growth (Fp-Growth) on Outdoor Product Sales Data,” *Int. J. Educ. Res. Soc. Sci.*, vol. 2, no. 1, pp. 132–139, 2021, doi: 10.51601/ijersc.v2i1.45.

[7] K. Kaur and H. K. Saini, “Retailer’s awareness regarding use of visual merchandising in apparel retail stores,” *Int. J. Home Sci.*, vol. 8, no. 1, pp. 103–108, 2022.

[8] V. P. S. Arora, S. Sharma, and N. K. Singh, “Nature and Functioning of Visual Merchandising in Organized Food Retailing,” *Agric. Econ. Res. Rev.*, vol. 20, no. June, pp. 441–456, 2007.

[9] M. B. R. Vaja, "Retail management," *Int. J. Res. Anal. Rev.*, vol. 2, no. 1, pp. 22–28, 2015, doi: 10.4324/9781315560854-16.

[10] I. Surjandari and A. C. Seruni, "Design of Product Placement Layout in Retail Shop Using Market Basket Analysis," *MAKARA Technol. Ser.*, vol. 9, no. 2, pp. 43–47, 2010, doi: 10.7454/mst.v9i2.379.

[11] F. R. Feijó and D. Botelho, "Effect of factors related to merchandising on retail sales," *RAE Rev. Adm. Empres.*, vol. 52, no. 6, pp. 628–642, 2012, doi: 10.1590/S0034-75902012000600005.

[12] H. Chaniago, "Investigation of Factors Influencing Traditional Retail Success in Small Cities in Indonesia," *J. Appl. Econ. Sci.*, vol. 1, no. 67, pp. 65–75, 2020, doi: 10.14505/jaes.v15.1(67).05.

[13] I. Sachdeva and S. Goel, "Retail store environment and customer experience: a paradigm," *J. Fash. Mark. Manag.*, vol. 19, no. 3, pp. 290–298, 2015, doi: 10.1108/JFMM-03-2015-0021.

[14] D. Law, C. Wong, and J. Yip, "How does visual merchandising affect consumer affective response?: An intimate apparel experience," *Eur. J. Mark.*, vol. 46, no. 1, pp. 112–133, 2012, doi: 10.1108/03090561211189266.

[15] P. Singh, N. Katiyar, and G. Verma, "Retail Shoppability : The Impact Of Store Atmospherics & Store Layout On Consumer Buying Patterns," vol. 3, no. 8, pp. 15–23, 2014.

[16] G. I. Cordova, S. Del Rosario Guerreros Surichaqui, F. Del Pilar Palomino Ricaldi, and W. Vicente-Ramos, "Impact of visual merchandising on the purchase decision of consumers from retail stores in central Peru," *Manag. Sci. Lett.*, vol. 10, no. 11, pp. 2447–2454, 2020, doi: 10.5267/j.msl.2020.4.005.

[17] A. Pratama, B. Mulyawan, and T. Sutrisno, "Implementation of improved apriori algorithm for drug planograms in XYZ pharmacies and analysis of medicinal procurement planning using ABC critical index methods," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1007, no. 1, pp. 1–7, 2020, doi: 10.1088/1757-899X/1007/1/012186.

[18] A. Fachrurozi, M. Junaedi, J. L. Putra, and W. Gata, "Algorithm Implementation Of Interest Buy Apriori Data On Consumer Retail Sales In Industry," *J. Informatics Telecommun. Eng.*, vol. 4, no. 1, pp. 48–56, 2020, doi: 10.31289/jite.v4i1.3775.

[19] H. Kusumo, D. Marlina, M. Novita, and M. T. Anwar, "Analysis of transaction patterns at drug store with Apriori Algorithm," *J. Phys. Conf. Ser.*, vol. 1869, no. 1, pp. 1–5, 2021, doi: 10.1088/1742-6596/1869/1/012070.

[20] T. A. Kumbhare and P. S. V Chobe, "An Overview of Association Rule Mining Algorithms," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 1, pp. 927–930, 2014.

[21] C. N. S. Kumar, V. S. Ramulu, K. S. Reddy, S. Kotha, and C. M. Kumar, "Spatial Data Mining using Cluster Analysis," *Int. J. Comput. Sci. Inf. Technol.*, vol. 4, no. 4, pp. 71–77, 2012, doi: 10.5121/ijcsit.2012.4407

[22] Z. Syahputri, S. Sutarman, and M. A. P. Siregar, "Determining The Optimal Number of K-Means Clusters Using The Calinski Harabasz Index and Krzanowski and Lai Index Methods for Groupsing Flood Prone Areas In North Sumatra," *Sinkron*, vol. 9, no. 1, pp. 571–580, 2024, doi: 10.33395/sinkron.v9i1.13246.

[23] T. S. Madhulatha, "An overview of clustering methods," *IOSR J. Eng.*, vol. 2, no. 4, pp. 719–725, 2012, doi: 10.3233/ida-2007-11602.

[24] N. Riyadi, M. F. Mulki, and R. Susanto, "Analysis of Customers Purchase Patterns of E-Commmerce Transactions Using Apriori Algoritm and Sales Forecasting Analysis With Weighted Moving Average (WMA) Methods," *Sci. Res. J.*, vol. VII, no. VII, pp. 45–58, 2019, doi: 10.31364/scirj/v7.i7.2019.p0719670.

[25] A. P. Windarto, "Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method," *Int. J. Artif. Intell. Res.*, vol. 1, no. 2, pp. 26–34, 2017, doi: 10.29099/ijair.v1i2.17.

[26] Y. Hefer and E. C. Nell, "Visual merchandising displays: the fashion retailer's competitive edge," *J. Gov. Regul.*, vol. 4, no. 4, pp. 408–411, 2015.

[27] A. R. Biscaia, M. J. Rosa, P. M. e Sá, and C. S. Sarrico, "Assessing customer satisfaction and loyalty in the retail sector," *Int. J. Qual. Reliab. Manag.*, vol. 34, no. 9, pp. 1508–1529, 2017.

[28] T. Scheffer, "Finding association rules that trade support optimally against confidence," *Intell. Data Anal.*, vol. 9, no. 4, pp. 381–395, 2005, doi: 10.3233/ida-2005-9405.

[29] M. R. Ridwan and H. Retnawati, "Application of Cluster Analysis Using Agglomerative Method," *Numer. J. Mat. dan Pendidik. Mat.*, vol. 5, no. 1, pp. 33–48, 2021, doi: 10.25217/numerical.v5i1.1538.

[30] T. Kyriazos and M. Poga, "Dealing with Multicollinearity in Factor Analysis: The Problem, Detections, and Solutions," *Open J. Stat.*, vol. 13, no. 03, pp. 404–424, 2023, doi: 10.4236/ojs.2023.133020.

[31] B. M. Prakash and Mishra Vivek, "Impact of Store Location and Layout on Consumer Purchase Behavior in Organized Retail," *Anvesha*, vol. 10, no. 1, pp. 10–21, 2017.

[32] I. A. Castro, A. C. Morales, and S. M. Nowlis, "The influence of disorganized shelf displays and limited product quantity on consumer purchase," *J. Mark.*, vol. 77, no. 4, pp. 118–133, 2013, doi: 10.1509/jm.11.0495.

[33] K. R. Shahapure and C. Nicholas, "Cluster Quality Analysis Using Silhouette Score," *2020 IEEE 7th Int. Conf. data Sci. Adv. Anal.*, vol. 1, no. 1, pp. 747–748, 2020.

[34] M. A. Sulaiman and M. N. Ahmed, "The essential elements of organized retail stores in influencing customers to stores," *Int. J. Appl. Bus. Econ. Res.*, vol. 15, no. 6, pp. 1–6, 2017, doi: 10.13140/RG.2.2.30709.12004.

[35] V. Jain, A. Sharma, and P. Narwal, "Impact of Visual Merchandising on Consumer Behaviour towards Women's Apparel," *Int. J. Res. Manag.*, vol. 5, no. 2, pp. 106–117, 2013.

[36] S. W. L. Fong, T. P. Kian, Y. S. Fern, and W. Vincent, "Monetary and Image Influences on the Purchase Decision of Private Label Products in Malaysia," *J. Adv. Manag. Sci.*, vol. 3, no. December, pp. 312–318, 2015, doi: 10.12720/joams.3.4.312-318

[37] N. Haj-Salem, J. C. Chebat, R. Michon, and S. Oliveira, "Why male and female shoppers do not see mall loyalty through the same lens? The mediating role of self-congruity," *J. Bus. Res.*, vol. 69, no. 3, pp. 1219–1227, 2016, doi: 10.1016/j.jbusres.2015.09.011.

[38] J. Leonard, "The Relationship between Color Preference and Consumer Products," *Undergrad. Psychol. Res. Methods J.*, vol. 1, no. 15, pp. 56–97, 2013.

[39] J. Clement, T. Kristensen, and K. Grønhaug, "Understanding consumers' in-store visual perception: The influence of package design features on visual attention," *J. Retail. Consum. Serv.*, vol. 20, no. 2, pp. 234–239, 2013, doi: 10.1016/j.jretconser.2013.01.003.

[40] R. Chandrashekaran and R. Suri, "Effects of gender and price knowledge on offer evaluation and channel transition in retail and e-tail environments," *J. Prod. Brand Manag.*, vol. 21, no. 3, pp. 215–225, 2012, doi: 10.1108/10610421211228838.

[41] A. Hübner, F. Schäfer, and K. N. Schaal, "Maximizing Profit via Assortment and Shelf-Space Optimization for Two-Dimensional Shelves," *Prod. Oper. Manag.*, vol. 29, no. 3, pp. 547–570, 2020, doi: 10.1111/poms.13111.

[42] I. Štulec, K. Petljak, and A. Kukor, "The Role of Store Layout and Visual Merchandising in Food Retailing The synchronized swimming club Medveščak," *Eur. J. Econ. Bus. Stud.*, vol. 2, no. 1, pp. 139–152, 2016.

[43] H. Kauppinen-Räisänen, "Strategic Use of Colour in Brand Packaging," *Packag. Technol. Sci.*, vol. 27, no. 8, pp. 663–676, 2014, doi: 10.1002/pts.

[44] B. S. Wijaya, M. S. Shihab, S. A. Wijaya, D. Rudianto, and A. Sugandi, "In-Store Brand Communication: When Shelf-Space and Display Seduce Consumers," *Humanit. Soc. Sci. Rev.*, vol. 8, no. 4, pp. 984–998, 2020, doi: 10.18510/hssr.2020.8495.