khazanah informatika
Jurnal Ilmu Komputer dan Informatika

Enhancing Retail Store Layout for Impulsive Buying Using Market Basket Analysis and the Apriori Algorithm

Sri Widiyanesti [1*], Muhammad Naufal Hakim [1], Nur Syamsiyah [2], Yogia Mugi Dirohmat [1]
*correspondence: Widiyanesti@telkomuniversity.ac.id
[1]Economy and Business Department
Telkom University
Bandung
[2]Agriculture Department
Universitas Padjajaran
Sumedang

Abstract-Retail serves as a crucial element in connecting product to end-customer. Accordingly, product assortment and placement are key factors in enhancing a store's attractiveness and promote convenience shopping. Therefore, customizing retail store layout must abide with customer behaviour. Market basket analysis (MBA) and association rule is the common framework to understand customer behaviour through historical transaction data. Yet, it can be extended to inform store layout improvement based on buying patterns. The current study aims to unveil customer buying pattern through MBA and association rules, then, use the collected insights to propose a new store layout design. We employed the Apriori algorithm to extract itemset relationships from the historical transaction data of a local convenience store brand. Furthermore, we integrate leverage metric to strengthen rule validation, offering more reliable interpretation compared to prior studies. Our findings suggest five solid rules that became the foundation of the proposed store layout, including a notably strong relationship between snack and drink products. The proposed framework can be adopted by retail businesses to improve store layout design tailored to their customer buying pattern.

Keywords: market basket analysis, association rules, layout planning, retail

Article info: *submitted January 31, 2024, revised May 27, 2025, accepted June 2, 2025*

1. Introduction

Retail stores represent the final point of a distribution chain, serving as the main interface between products and end-customer. As a brick-and-mortar establishments, they are typically associated with a big building and vast product arrangements. However, nowadays, retails also came in smaller size such as convenience store. Its smaller format allows businesses to get closer towards their customer by locating itself in a residential area proximity and increase area attractiveness [1]. To broaden the perspective, convenience store chain existence in small towns and rural areas contributes to consumer behaviour change and perceived as an extension of big cities influence [2]. In contrast to its advantages, there is also challenges attributed to small retails including optimal product assortment, space utilization, and effective information management system [3, 4]. Accordingly, we can infer that the mentioned issues can be addressed by implementing effective store layout strategy that covers adequate product variant and placement to drive sales. The effective layout strategy refers to store layout that accommodate shopping convenience based on customer behaviour. This was highlighted by prior research on optimizing store layout using historical transaction data [5, 6, 7, 8, 9]

In Indonesia, small retail convenience store thrives both in big cities and smaller towns. This market segment is predominantly dominated by two populous convenience store chains: Alfamart and Indomaret. As of 2021, there were approximately 19.000 Indomaret stores and 16.000 Alfamart stores operating across the country [10]. An interesting strategy employed by both brands is frequent store placement in close proximity to one another. Solnet et. al., [11] argue that this is a strategic move that allows subsequent store developers to save time and resources by reducing the need of extensive market research.

As mentioned earlier, convenience chains share a common challenge: 'how to manage a store's layout to enhance customer's shopping activity?'. The challenge seems not complicated as it looks since when two stores provide the same products, then logically, they can copy each other layout. However, Rizki & Supriadi [12] noted that Alfamart is known for its responsiveness, while Indomaret emphasizes stability. This means that both brands are practically employed similar store layout yet present different feels for customers. Hence, simply copying each other is not a practical solution since the consumers possess a brand image of the store and memorize it as a leisurely feeling as they shop. Therefore, a handful of resolution in terms of layout needs to be customized and updated periodically based on the store's consumers' behaviour. Market basket analysis is the commonly used method to gain insights into consumers' shopping behaviour by mapping consumers' choices [6, 9]. However, it is not exclusive for retail store since studies have proven the use of market basket analysis on different services, such as libraries [13], amusement arcades [14], even hotels [15]. The mentioned studies show the benefit of market basket analysis which will be discussed further in the current study.

Market basket analysis is possible using algorithms in association rule data mining. In the beginning, market basket analysis (MBA) was used by retailers to analyze the content of consumers' *baskets* which is their shopping items. Nevertheless, businesses catch the opportunities to use MBA to suggest a combination of products for special promotions, give the insight to brand loyalty or co-branding, and even build a more effective store layout [13]. We consider the type of retailers that may gain an advantage with the MBA are supermarkets, convenience stores, and fashion stores. In a bigger store format such as supermarket, Kawengian et. al., [5] highlight the importance of combining MBA with two other approaches: OCVR and activity chart to thoroughly understand the store's customer journey. However, we argue that it is not resource efficient if implemented in the smaller store such as convenience store. To broaden the perspective, Setiawan & Mulyanti [8] implemented MBA in digital platform to design fashion product recommendation systems on a retail store's website by unveiling customer's buying patterns based on historical transactions data. We concur that this mechanism is best for remarketing process. Santoso [6] performed MBA to an Indonesia's local brand convenience store branches and prevailed in obtaining 22 sets of rules that can enhance the store's sales strategy by rearranging product display based on the extracted rules. Furthermore, MBA can help retails to bundle products based on the extracted association rules and placed them in high traffic section or close to high margin products to increase impulsive buying possibility [9]. To the very end, MBA implementation was aimed to improve shopping experience through store layout customization [7, 16]. The current study also aims to obtain similar rules, but we performed MBA using a different software compared to Hamdani's study. In addition, we present the updated store layout using based on the MBA findings similar to Surjandari & Seruni [7].

One of the keys to success in retailing is understanding the consumers' behaviour. In this digital era, retailers can gain insights from the historical transaction data built upon consumer activities. Seidou [17] highlighted customer's purchasing decision in convenience store is influenced the following factors: store format (i.e., location, product assortment, service, price, and atmosphere) and sales promotion (i.e., free sample, price reduction, shopper card, and coupons). Pallikara et. al., [18] argue on distinct factors that influence impulse buying behaviour into two groups; first is the 'situational factor'; money and time availability, family or social Influence, store environment, credit card availability, and momentary mood. The second is the 'external stimuli factor': in-store promotion, offers/discounts, bonus packs, large merchandise, product placement, peer influence, in-store service, and friendly employees. The current study attempt to use the external stimuli factors to construct store layout to enhance the possibility of impulsive buying behaviour.

There are two common algorithms that can be used to extract association rules in data mining: Apriori and Frequent Pattern Growth (FP-Growth). The FP-Growth algorithm is an improved version of its predecessor, the Apriori algorithm. Patil & Patil [19] study proved that the FP-Growth algorithm outperformed the apriori algorithm regarding the time and memory required to process the dataset. However, Apnena & Firdhani [20] argue that the apriori algorithm is better than the FP-Growth algorithm in the association rules formed to process sales transactional data. We took this opinion as justification for why we performed the current study using the Apriori algorithm instead of the improved version of it (FP-Growth algorithm).

The current study employs objective measures to extract insights from association rules. Tan et. al., [21] highlight and summarize up to 21 objective metrics that can be used to analyze and interpret association rules. Despite the vast and diverse measures, previous studies [6, 5, 7, 16, 9] predominantly rely on three key metrics: support, confidence, and lift. Our study adopts these commonly used three metrics and further incorporate leverage to enhance the quality of the extracted insights. Leverage metric offers an absolute measure on an association rule strength and is robust to rare itemset compared to lift metric [21]. Thereby, including this metric provides a more reliable basis for optimizing the store's new layout compared to approaches in prior research. Additionally, we apply a confidence threshold of 100%, in contrast to previous studies that use lower thresholds i.e., 50% or 90% [16, 9, 7]. The description of each metric is provided in table 1.

The contribution of this study is an improvement to the store layout for the populous convenience store brand in Indonesia. First, we collected data on the current layout and sales transaction data for a given period. Second, we perform association-based data mining on a respectable data mining application to determine consumers buying patterns through rules. Lastly, we designed the brand-new store layout from the findings.

2. Methods

The current study's methodology uses the apriori algorithm to perform a market basket analysis of the given sales transaction data. We employ the modified research procedure from [6]. We perform a literature review to study and compare how previous studies perform market basket analysis, thus identifying the research gap. Then, we visit the store to observe its current layout. Next, we collected the sales transaction data for July 2022 and August 2022; the total amount of the collected data is no less than 300 transactions. The data undergo pre-processing and are transformed into a corpus before the author performs association-based data mining using the Apriori algorithm and evaluates the results with Orange, the data mining software. Based on the outcome, we designed a new store layout based on the dominance of certain products and the relationships between them. Therefore, it should increase the comfort level of the store's consumers and encourage impulsive buying behaviour.

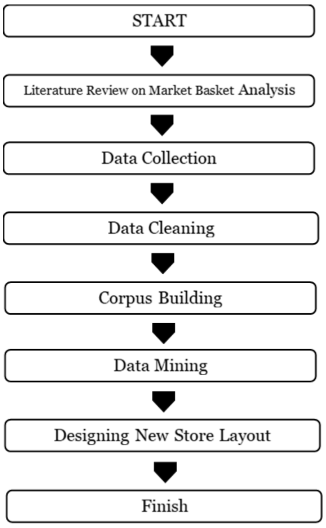

Figure 1. Research Workflow

1. The Dataset

The data used in the current study is sales transaction data for July and August in 2022; the data format can be seen in table 1. During the data collection process, the author briefly interviewed the store manager. The findings are that the store rarely changes its layout since the layout planning is centralized; furthermore, the central command rarely asks to do so. The manager is aware that the placement of some products is less effective in promoting them.

Table 1. Dataset Sample

Time	Transaction ID	Items	Qty	Price
17:37:30	V8532210011 108Y387	Kiranti 250ml	2	Rp6.400
17:37:30	V8532210011 108Y387	Sampoerna Mild	1	Rp25.700
19:07:44	V8932210011 108Y387 Rp9.900	Tao Kenoi Rumput Laut 15g	1	Rp9.900
19:07:44	V8932210011 108Y387	Pringles Original 107g	1	Rp23.800
19:07:46	V8932210011 108Y387	Signature 12	1	Rp18.500

The acquired data is cleaned by removing attributes that are not necessary, such as the time, transaction id, and price. The quantity attributes are presented in the transformed dataset as a repetition of the same product in each transaction. Furthermore, the author reduced the number of items in each transaction to only seven unique products, according to the mean of the products in the whole transaction. Therefore, the cleaned data is formatted as in table 2 and transformed into text format (.csv) before entering the Orange environment.

Table 2. Transformed Dataset Sample

Item				
Kiranti	Sampoerna Mild	-	-	-
Tao Kenoi Rumput Laut	Pringles Original	Signature	Good Day	-

3. Result

In this section, the author presented a step-by-step of analyzation process using the Orange data mining application. We perform two stages of association rules with different settings of the algorithm in between.

The analysis workflow is visualized in Figure 1. The loaded data is processed using the association rules feature; the author pre-defined the minimum support as 0.08 and the confidence as 1.0. The result is 14 rules that met the conditions as visualized in table 3.

Table 3. 1st Association Rules Extraction

Itemset (antecedent > consequent)	Support	Confidence
Shampoo Sunsilk > Nuvo	0.059	1.000
Sunlight > Rinso	0.059	1.000
Sunlight > Nuvo	0.059	1.000
Teh Pucuk > The kotak	0.065	1.000
Taro > Pringles Original	0.065	1.000
Pringles Original > Pringles Black Papper	0.065	1.000
Tao Kenoi Rumput Laut > Pringles Original	0.065	1.000
Silverqueen > Biskuit Roma	0.068	1.000
Silverqueen > Ice Cream Walls	0.068	1.000
Pocari Sweat > UC 1000	0.080	1.000
Im Boost > UC 1000	0.080	1.000
Redoxon > Im Boost	0.080	1.000
Piattos > Le Minerale	0.083	1.000
Jappota > Lays	0.083	1.000

Interpretation for the first and third rules are based on the given data, consumers in this store are more likely to purchase 'Nuvo' products alongside 'Shampoo Sunsilk' or 'Sunlight'. The interpretation is the same for the other rules according to the name of the products in each rule. Since the confidence value is '1.0' then; we can infer that the consequent is 100% likelihood to be bought along with the first picked item (antecedent). The author adds more measurements in the second analysis since confidence cannot detect statistical dependencies [22].

The next step for the analysis is to change the number of antecedents to three; this means the algorithm will present the fourth item added to the basket based on the previous three items. The author also accounts for the value of lift and leverage to add more insights into the resulting rules. The result of the second analysis is fiver rules that are visualized in the Table 4.

Table 4. 2nd Associaton Rule Extraction

Itemset (antecedent > consequent)	Support	Confidence	Lift	Leverage
Shampoo Sunsilk, Nuvo, Rinso > Sunlight	0.059	1.000	16.850	0.056
Tao Kenoi, Teh kotak, Taro > Teh Pucuk	0.065	1.000	15.318	0.061
Pringles Original, Pringles Blackpapper, Tao Kenoi > Teh Kotak	0.065	1.000	15.318	0.061
Imboost, Redoxon, Pocari Sweat > UC1000	0.080	1.000	14.433	0.070
Piattos, Lays, Japotta > Le Minerale	0.083	1.000	12.036	0.076

There are fewer rules compared to the first analysis though the minimum support and confidence have stayed the same. The rising antecedent plays the most significant factor in the resulting rules. In compliance with the fifth rule, consumers are more likely to buy 'Le Minerale' along with 'Piattos', 'Lays', and 'Japotta'. Although the confidence value is the same for each rule, the interpretation now differs based on the lift and leverage measurements. The higher the lift or leverage, the more itemset in the presented rules are more likely to sell together. From both measurements, we can infer where to look for measurements when discerning the rules. The lift shows dependencies (strong association) on the low support rules, and the leverage is the exact opposite.

The exciting rules to be discussed are the second and third, which have the same value for the four measurements despite having different consequences. This means both rules can be used to categorize a particular consumer behaviour since both rules contain the same item category: snacks and drinks. The author assumed that the customer was girl's teenager. The rules encourage the author to design the proposed layout, where the store's consumers can quickly notice the snacks and drinks. The store manager might develop product bundling for special events or promotions with this knowledge. The results in the second analysis also served as the basis for designing the proposed store layout.

4. Discussion

The current section will present the proposed store layout and discuss how it is better than the current layout. We make product categorizations to make the reader easier to digest the discussed matter. Nadif & Vanany [23] presented product categorizations for retail business encompasses snacks, drinks, cooking ingredients, instant foods, health & beauty products, home products, frozen food, canned food, and bread & cake. Our study modifies that product categorization by adding two more categories, tobacco and fruit and vegetables. The current store layout is visualized in Figure 2, and it was confirmed during the interview with the store's manager that the layout rarely changes. Therefore, we validate the proposed store layout design since there are no changes to the current layout based on the given sales transaction data.

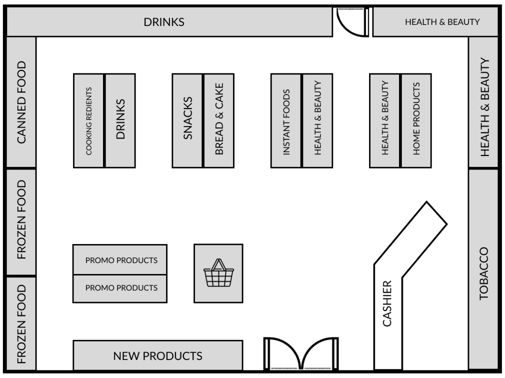

Figure 2. The Current Layout

The 'tobacco' was put behind the cashier for safety reasons, mainly to prevent under aged person to buy it and avoid shoplifting. The 'promo products' section was used to display the product bundling, discount products, and packed fruits or vegetables, while the 'new products' section was used to promote the new arrival products.

The improvement to the current store layout is visualized in Figure 3, serving as our proposed layout. We are optimistic that the design will enhance consumers' convenience during the shop since it was built using the extracted sales pattern from the sales transaction data. The highlight of changes in the proposed layout is that first, we put the 'instant foods' close to 'snacks', then we added the 'fruit & vegetable' product category, then placed it close to 'health & beauty'. Lastly, we swapped the 'bread & cake' section with 'Cooking Ingredients'.

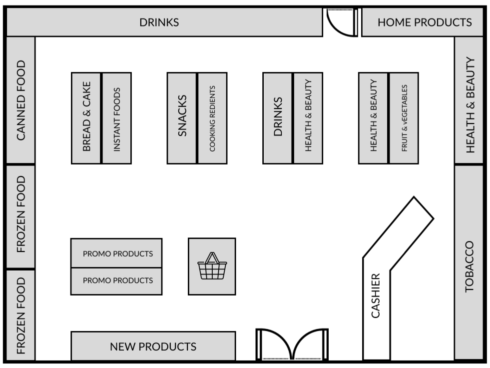

Figure 3. The Proposed Layout

The reasoning for changes in the proposed layout is based on the previously acquired association rules. Then, we placed the less interesting products close and centred around the highly interesting ones, allowing impulsive buying. The less interesting products are identified as products with low support value in the dataset, and the highly interesting products are the opposite. The addition of the ‘fruit & vegetables’ category is to avoid mixing them with other products, such as in the previous layout. Therefore, each product category can be recognized distinctively from one another.

The important notes on the layout updates are that they are dynamics. The new proposed layout is constructed using two-month historical data that explain consumers’ shopping behavior at the store. The layout may change or be updated periodically to ensure it follows their customer’s behaviour. Hence, the store will continuously encourage the occurrence of impulsive buying behaviour. Nevertheless, the proposed store layout can be different depends on the employed algorithm e.g., FP-Growth and HUIM which covers an item’s profit margin that are essential in maximizing store’s profitability [24, 25].

5. Conclusion

The current study proposed a brand-new store layout through MBA using association rules and the Apriori algorithm by analyzing historical sales transaction data between July and August 2022 from a branch store of the famous Indonesian local convenience store brand. Based on the extracted rules, it shows that the majority of customer loves to buy snacks and drinks as a bundle, hence, the store layout strategy should revolve around this high demand product categories to increase impulsive buying probability. Other retails can adopt our research framework to improve the store layout based on their customer buying pattern.

Our study contributes to the literature by introducing leverage metric to solidify the extracted association rule interpretation. Future research should explore other metrics to verify its suitability in understanding customer buying pattern to improve a store layout. In addition, comparative study can be done to identify the discrepancies between store layout suggested by different algorithm e.g., FP-Growth and HUIM. Accordingly, future studies can cover cost & benefit analysis and store’s profitability which factor the importance of having a strategic retail’s layout arrangement.

There are several limitations that can be addressed in future research. First, we use the Apriori algorithm, which, despite its simplicity, costs time and memory [19]. Second, the current study’s dataset is relatively small compared to prior research. This can be addressed in future research by expanding the data volume (more transaction over longer period). In addition, the data can be collected from several retail stores over a defined area/region to enhance future study’s scope of the research in attempt to understand larger population’s buying pattern. Lastly, our study did not account for item’s input sequence in the customer’s basket, thus, the extracted rules may not explain in-store customers’ journey.

In conclusion, our study lays the groundwork for a more strategic and data-driven approach to retail layout design. By moving beyond conventional evaluation metrics which promote better rule validation. While future research is needed to scale and refine this framework, the practical implications are clear: a well-informed layout aligned with actual customer buying pattern can be a powerful tool to improve in-store experience.

Acknowledgement

-

Reference

[1] M.-L. Allain and A. Épaulard, "Small Retail Businesses: Decline or Mutation?," Notes du conseil d’analyse économique, vol. 77, no. 2, pp. 1-12, 2023.

[2] M. Twardzik and B. Kucharska, "The Impact of Retail Chains on the Structure of Small Towns and Rural Areas in Poland," in 5th World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium – WMCAUS, Prague, 2020.

[3] Y. Gu and Y. Wu, "Highlighting supply-abundance increases attraction to small-assortment retailers," Journal of Retailing, vol. 99, no. 3, pp. 420-439, 2023.

[4] W. Gunawardana and N. Wedasinghe, "Effective Use of Retail Store Management System for Small Retail Store," in 12th edition of International Conference on Business and Information (ICBI 2021), Colombo, 2021.

[5] A. R. Kawengian, I. H. Lahay and J. D. Giu, "Redesain Tata Letak Produk Berdasarkan Market Basket Analysis," JISI: Jurnal Integrasi Sistem Industri, vol. 12, no. 1, pp. 49-58, 2025.

[6] M. H. Santoso, "Application of Association Rule Method Using Apriori Algorithm to Find Sales Patterns Case Study of Indomaret Tanjung Anom," Brilliance: Research of Artificial Intelligence, vol. 1, no. 2, pp. 54-66, 2021.

[7] I. Surjandari and A. C. Seruni, "Design of Product Placement Layout In Retail Shop Using Market Basket Analysis," Makara Journal of Technology, vol. 9, no. 2, pp. 43-47, 2005.

[8] A. Setiawan and R. Mulyanti, "Market Basket Analysis dengan Algoritma Apriori pada Ecommerce Toko Busana Muslim Trendy," JUITA: Jurnal Informatika, vol. 8, no. 1, pp. 11-18, 2020.

[9] T. Joe, S. R and K. Sekar, "Optimization of Store Layout using Market Basket Analysis," International Journal of Recent Technology and Engineering, vol. 8, no. 2, pp. 1459-1463, 2019.

[10] H. Nurhayati-Wolff, "Number of stores of leading convenience store chains in Indonesia in 2021," 22 September 2022. [Online]. Available: https://www.statista.com/statistics/1228079/indonesia-number-of-stores-of-convenience-stores/. [Accessed 12 December 2022].

[11] A. Anbalagan, "Unique battle between international and local convenience stores in Indonesia," 26 October 2018. [Online]. Available: https://thelowdown.momentum.asia/unique-battle-between-international-and-local-convenience-stores-in-indonesia/. [Accessed 14 December 2022].

[12] D. F. Rizki and N. Supriadi, "Comparison of Alfamart and Indomaret Supply Chain Efficiency in Handling Consumer Demand in Indonesia," American Journal of Economic and Management Business, vol. 3, no. 6, pp. 144-153, 2024.

[13] S. J. Cunningham and E. Frank, "Market basket analysis of library circulation data," in 6th International Conference on Neural Information Processing, Perth, 1999.

[14] S. Halim, T. Octavia and C. Alianto, "Designing facility layout of an amusement arcade using market basket analysis," Procedia Computer Science, vol. 161, pp. 623-629, 2019.

[15] D. Solnet, Y. Boztug and S. Dolnicar, "An untapped gold mine? Exploring the potential of market basket analysis to grow hotel revenue," International Journal of Hospitality Management, vol. 56, pp. 119-125, 2016.

[16] D. Alfitra, M. A. M. Afdal, M. Fronita and E. Saputra, "Market Basket Analysis for Determine Goods Layout Using FP-Growth

Algorithm," Sistemasi:Jurnal Sistem Informasi, vol. 13, no. 4, pp. 1651-1661, 2024.

[17] S. Hafissou, "The impact of store formats and sales promotion towards consumer's purchase decision: Case study of Indomaret in Bandung city," Journal of Administrative and Business Studies, vol. 6, no. 5, pp. 164-175, 2020.

[18] V. Pallikkara, P. Pinto, I. T. Hawaldar and S. Pinto, "Impulse buying behaviour at the retail checkout: An investigation of select antecedents," Business: Theory and Practice, vol. 22, no. 1, pp. 69-79, 2021.

[19] M. Patil and T. Patil, "Apriori Algorithm against Fp Growth Algorithm: A Comparative Study of Data Mining Algorithms," SSRN Electronic Journal, 2022.

[20] R. D. Apnena and F. Faujiyah, "Comparison of Apriori and FP-Growth Algorithms in Recommending Promotional Itemsets for Entrepreneurs Using the Lift Ratio Method (Case Study : Berkah Jaya, Elektrik 67, E-Commerce," Application of Information Technology (APIT), vol. 1, no. 1, pp. 61-71, 2022.

[21] P. N. Tan, V. Kumar and J. Srivastava, "Selecting the right objective measure for association analysis," Information Systems, vol. 29, no. 4, pp. 293-313, 2004.

[22] F. Berzal, I. Blanco, D. Sanchez and M. A. Vila, "Measuring the accuracy and interest of association rules: A new framework," Intelligent Data Analysis, vol. 6, no. 3, pp. 221-235, 2002.

[23] N. Nadif and I. Vanany, "Analisa Disrupsi Permintaan pada Ketersediaan Produk di Industri Retail Nuarania," Prozima, vol. 5, no. 1, pp. 12-20, 2021.

[24] W. A. Kamakura, "Sequential market basket analysis," Marketing Letters, vol. 23, no. 3, pp. 505-516, 2012.

[25] H. Yao and H. J. Hamilton, "Mining itemset utilities from transaction databases," Data & Knowledge Engineering, vol. 59, no. 3, pp. 603-626, 2006.