Optimizing the Grocery Store Layout Using Market Basket Analysis

Hellenita Chiani[a], Genrawan Hoendarto [b]

[a]*Department of Digital Business, Faculty of Technology Information Technology, Widya Dharma Pontianak University, Pontianak, Indonesia 78117*
hellenitachiani24@email.com
[b]*Department of Informatics, Faculty of Technology Information Technology, Widya Dharma Pontianak University, Pontianak, Indonesia 78117*
genrawan@widyadharma.ac.id

First Author Biography
Hellenita Chiani was born in Pontianak, West Borneo, Indonesia, in 2004. She is presently enrolled in Widya Dharma Pontianak University's undergraduate program in digital business. Her research interests include machine learning-based big data processing, especially its use in business analytics.

Second Author Biography
Genrawan Hoendarto was born in Pontianak, West Borneo, Indonesia, in 1968. He completed his bachelor's degree in electrical engineering at Tanjungpura University, Pontianak, Indonesia, and his master's degree in computer science at Eresha School of Management and Computer Science, Jakarta, Indonesia. He is currently completing his doctoral degree in computer science at Sepuluh Nopember Institute of Technology Surabaya, Indonesia. His research interests include Machine Learning, the Internet of Things (IoT), and Building Automation.

Optimizing the Grocery Store Layout Using Market Basket Analysis

1. Abstract

The importance of store layout can affect customer shopping experience and store operational efficiency. This study uses the Market Basket Analysis (MBA) method to analyze customer purchasing patterns, identify relationships between product categories purchased together, and generate recommendations for developing a more effective store layout. Achan Store is a wholesale and retail store of groceries and snacks in Pontianak, West Kalimantan, Indonesia. This study uses quantitative methods and primary data in the form of sales transaction history. The population of this study is transaction data for 6 months from November 1, 2023 to March 30, 2024, with a sample of 1,203 transactions. The results of the MBA algorithm are evaluated based on the value of support, confidence, and lift. The results of this analysis indicate that product categories should be placed close together because they have close support and confidence values. For example, if the toiletries category shelves are placed adjacent to the bottle drinks shelf, then the toiletries shelf should be adjacent to the detergent shelf because they are more often purchased together. The resulting layout recommendations are expected to increase store efficiency and sales and improve customer shopping experience.

2. Introduction and Research Problem

Product layout influences customers' purchasing patterns and decisions that affect store performance (Hirpara & Parikh, 2021; Mowrey et al., 2018). It also influences the creation of a comfortable and effective shopping environment for customers (Valle et al., 2018), and usually customers come to the store with shopping records and start from the closest distance (Urban, 2022).This is because many customers are not only concerned with 'easy to use' products, but also with 'easy to find' aspects when shopping (Suryadi et al., 2022). Many decisions must be made for this layout, such as how much space or shelving is required for each product, where to place product categories relative to each other (Ozgormus & Smith, 2020), and ensuring that the resulting layout does not disrupt or impede in-store traffic and minimizes inefficient customer movement (Pantano et al., 2021).

The use of machine learning based on real data with Market basket Anylisis (MBA) is an effective tool for measuring or gaining knowledge about consumer shopping habits, patterns, and preferences (Liu et al., 2024). The use of MBA identifies which product categories are often purchased together (Griva et al., 2018). The transaction data collected in this study were evaluated using three important components: support, confidence, and lift. Support shows how popular the item is, confidence refers to how often item A is purchased with item B, for the lift value if it is above 1, then the purchase of product A together with B is not a coincidence or other meaning is more often purchased together (Mohd Rosli & Ibrahim Teo, 2022). Although each grocery store is different, the MBA method can generate preference closeness that is general and applicable.

Achan Store is a family business in the sense that it can be managed and controlled by the founder or other family members (Pereira et al., 2024) who have a specific role in achieving strategic objectives and sustainability in their markets (Rohemi et al., 2024). Currently, Achan Store does not yet have an efficient product layout. The placement of products per category is still random/separated, as well as several shelves that are deemed inappropriate placement and use so that they take up space

and make it difficult for customers when shopping. Therefore, the results of this study can be used as suggestions or references for Achan stores to maximize their product layout.

3. Methods

This quantitative research aims to obtain the best side-by-side positioning of products per category. Primary data comes from sales transaction data of the Achan Store. Achan Store became the research subject, which is a wholesale and retail store in Pontianak, West Kalimantan, Indonesia. 6 months of data from November 1, 2023 to April 30, 2024 serve as the research population with a sample of 1,203 transactions. The layout arrangement is considered the dependent variable, and the purchase pattern of the sales transaction is treated as the independent variable.

First, the sales data for 6 months with a total of approximately 1,500 products were separated into a new table consisting of product name, quantity purchased, total price, unit, transaction date, and receipt number. The product names were then grouped into 18 categories. Subsequently, a new table was created containing category columns, product name, transaction date, and receipt number. Furthermore, the data are sorted based on the receipt number. This data is used to create a binary table that contains a column of note numbers and the 18 categories specified, namely mineral water, snacks, instant drinks (bottles/boxes/cans), cigarettes, sachet drinks, instant noodles, kitchen spices, bread/biscuits/crackers, detergents, candy/chocolate, basic necessities (flour, oil, rice, sugar), pest cleaners, stationary, medicines, toiletries (personal hygiene products), and disposable products. The category column is filled with 1 if on the n-th note there is a transaction involving that product category, and will be filled with 0 if there is no purchase transaction.

This data table was then processed using Matlab software version R2020a. The processing results are used to identify customer purchase patterns by analyzing data from the most purchased product categories. This involves assessing the support value to determine the best-selling product categories as well as the confidence and lift values to establish a suitable layout of product categories placed close together and to assess data validation. The minimum support and confidence values are determined based on the average value of all data and other factors that reflect the reality of best-selling products.

4. Results and Findings

At this stage, data processing will use the Market Basket Analysis algorithm. The minimum values for support and confidence are set based on the average value of all categories. To determine customer buying patterns, the analysis focuses on 18 product categories from 1,203 transaction data for 6 months from November 1, 2023 to April 30, 2024.

The results of the analysis will reveal the most purchased product categories and the most frequently purchased pairs of categories together, such as the most purchased cigarette category with 87% support, the candy/chocolate category is suitable to be placed adjacent to snacks with 54% confidence and 1.64 lift). The results of this analysis will also help to rearrange the previously random product layout into a better layout based on the results of the confidence and lift calculations for each product category.

No.	Colors	Item color for each category
1		Cigarette
2		Detergent
3		Processed food/ready to eat
4		Baking Ingredients
5		Toiletries (makeup, skincare)
6		Pesticide
7		Medicine
8		Stationary
9		Snack
10		Sachets/Powder Drink (Instant Coffee)
11		Biscuits/Crackers
12		Bottles/Cans/Boxes Drink
13		Disposible Item (Cutlery/plastic bag)
14		Mineral Water
15		Grocery (Cooking Oil, Flour, Sugar, Rice)
16		Candy/Chocolate/Gummy
17		Instant Noodles
18		Seasonings/spicies
		Drafting Media for each Category
1		Table
2	1-34	Shelf
3		Wooden Pallets
4		Glass Display Case
5		Hanger
6		Werehouse

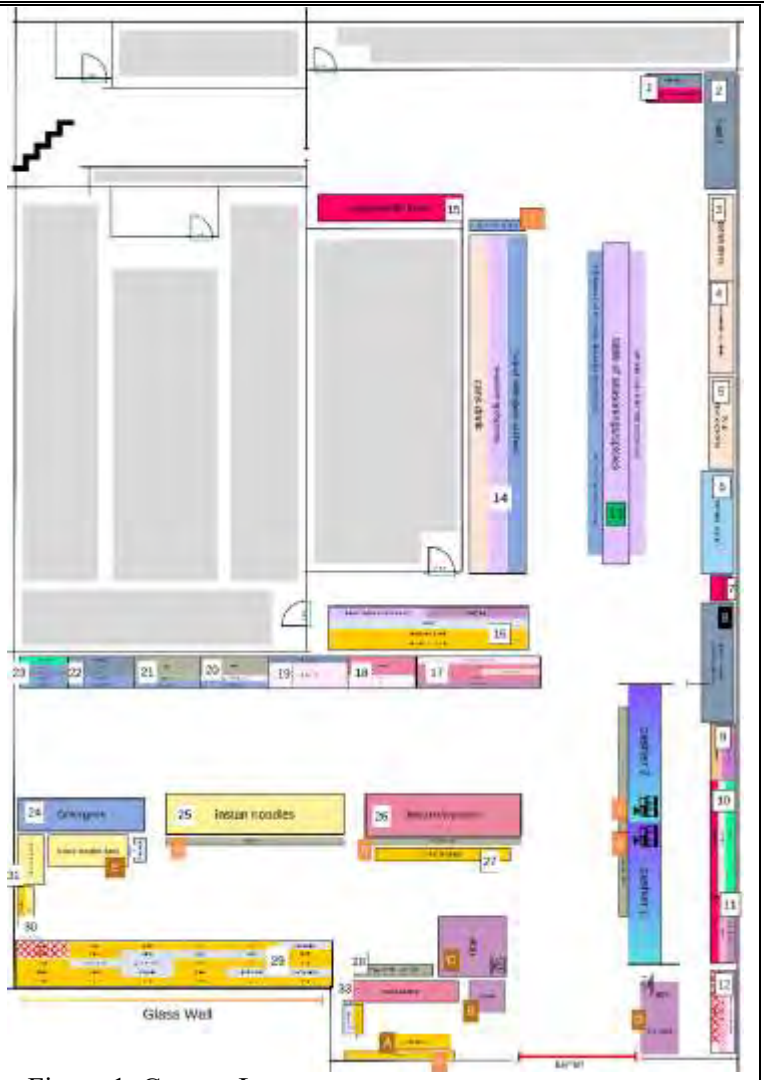

Figure 1. Current Layout

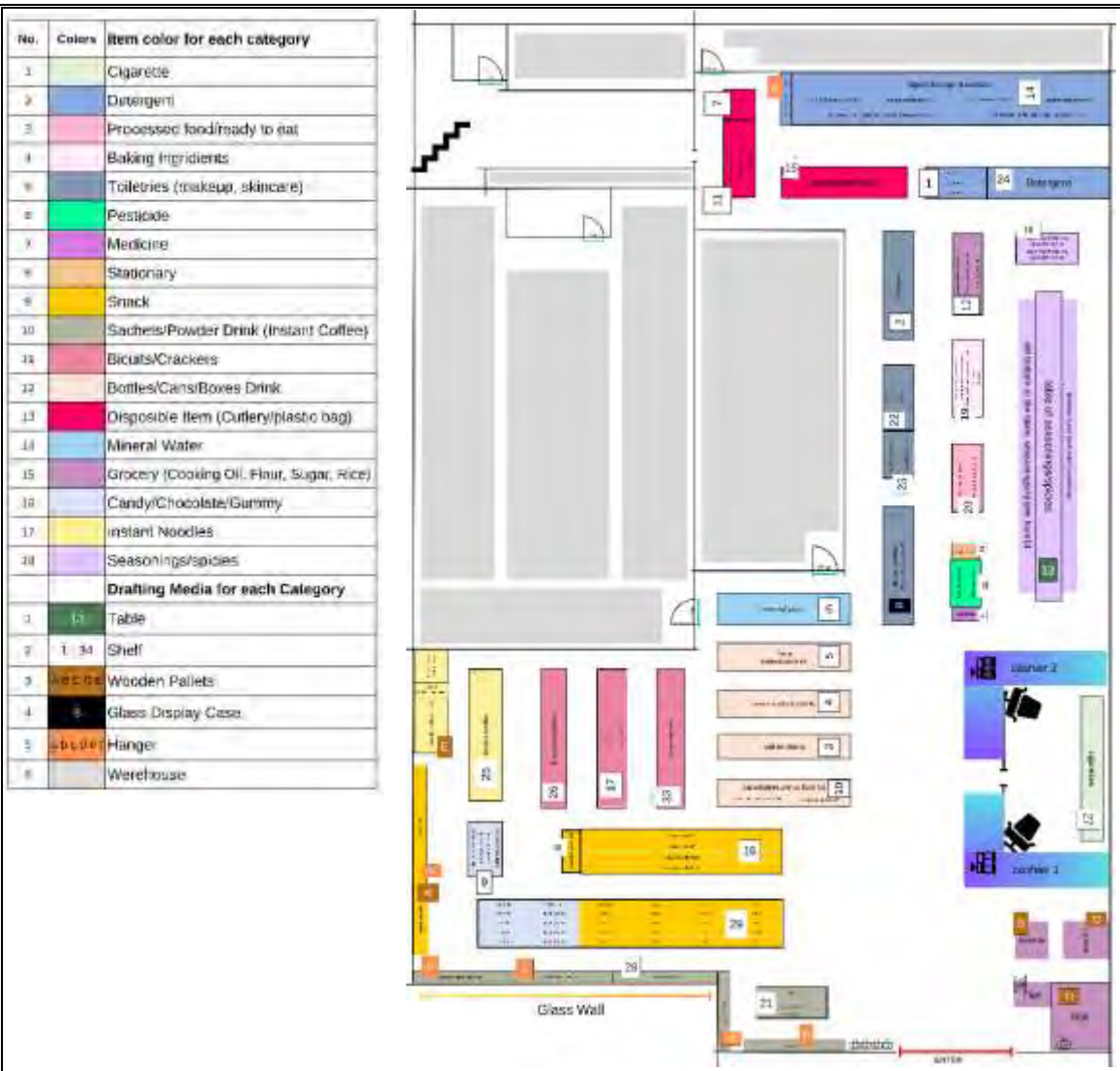

Figure 2. Recomendation Layout

The resulting layout proposal is a product relocation recommendation that can be implemented by Achan stores. Such as rearrangement on shelf number 29, the position of snacks and candies that were initially random to be structured, making it easier for customers and employees to find and organize products. Place snacks in the rightmost arrangement so that they are more easily visible to customers. Placing candy on the left because support for snacks (41%) is higher than for candy and chocolate (33%). Shelf number 29 is placed in the center because it has above-average confidence and lift for all categories except cigarettes. It has a positive relationship with all categories because it is often purchased together. The placement of instant noodles, which were originally placed on shelf number 30 and side by side with snacks, was moved to the wooden pallet (E). Because it is easier to reach and the shelf is of the right size. The original shelf is too small to accommodate all the instant noodles. Shelf number 30 can be allocated to other products. The sachet drink hanger (52% support) with the highest support is placed in the area that is most accessible and visible to customers. Reorganizing the cashier area into two sections was done to prevent long queues at one cashier at the time of payment.

5. Conclusions, Implications, and Significance

This study uses market basket analysis (MBA) to statistically generate data on customer purchasing patterns at Achan stores. The results of this analysis of customer purchasing patterns identify relationships between product categories that are purchased together. This relationship then becomes the basis for a more efficient layout.

The application of innovative data analysis methods such as MBA can be an effective tool for Achan stores to improve their business performance. The store owner can develop a more strategic and attractive layout. Store space can be optimized to improve product attractiveness and business performance efficiency. This also opens up opportunities to improve marketing strategies and product promotions that are more targeted.

The results of this study are significant in terms of efficiency and profitability in the grocery store business. The application of the MBA method in this study contributes to the development of best practices for managing grocery stores. The results of this study can be a foundation for grocery store owners to improve their business performance by organizing product layouts. In addition, it provides valuable insights for other researchers interested in data analysis in the field of retail management.

Reference

Griva, A., Bardaki, C., Pramatari, K., & Papakiriakopoulos, D. (2018). Retail business analytics: Customer visit segmentation using market basket data. *Expert Systems with Applications*, *100*, 1–16. https://doi.org/10.1016/j.eswa.2018.01.029

Hirpara, S., & Parikh, P. J. (2021). Retail facility layout considering shopper path. *Computers and Industrial Engineering*, *154*. https://doi.org/10.1016/j.cie.2020.106919

Liu, F., Huang, W., Zhang, J., & Fang, M. (2024). Corporate social responsibility in family business: Using machine learning to uncover who is doing good. *Technology in Society*, *76*. https://doi.org/10.1016/j.techsoc.2024.102453

Mohd Rosli, N. A. H., & Ibrahim Teo, N. H. (2022). MARKET BASKET ANALYSIS USING APRIORI ALGORITHM: GROCERY ITEMS RECOMMENDATION. *Advanced International Journal of Business, Entrepreneurship and SMEs*, *4*(14), 01–09. https://doi.org/10.35631/aijbes.414001

Mowrey, C. H., Parikh, P. J., & Gue, K. R. (2018). A model to optimize rack layout in a retail store. *European Journal of Operational Research*, *271*(3), 1100–1112. https://doi.org/10.1016/j.ejor.2018.05.062

Ozgormus, E., & Smith, A. E. (2020). A data-driven approach to grocery store block layout. *Computers and Industrial Engineering*, *139*. https://doi.org/10.1016/j.cie.2018.12.009

Pantano, E., Pizzi, G., Bilotta, E., & Pantano, P. (2021). Enhancing store layout decision with agent-based simulations of consumers' density. *Expert Systems with Applications*, *182*. https://doi.org/10.1016/j.eswa.2021.115231

Pereira, V., Laker, B., Bamel, U., Sharma, G. D., & Paul, H. (2024). Customer engagement strategies within family businesses in emerging economies: A multi-method study. *Journal of Business Research*, *174*. https://doi.org/10.1016/j.jbusres.2024.114508

Rohemi, Z. O., Stefania, G. C., Eduardo, M. C. A., Rafael, B. S. G., Lucia, R. V. A., Marina, R. O. A., Darío, R. H. J., & María, N. R. L. (2024). Clustering and Principal Component Analysis for the characterization of corporate governance in family businesses. *Procedia Computer Science*, *231*, 496–501. https://doi.org/10.1016/j.procs.2023.12.240

Suryadi, A., Cattleya, M., & Anissa, P. (2022). *Analysis of Data Mining at Supermarket X in Surabaya Using Market Basket Analysis to Determine Consumer Buying Patterns*. *2022*, 28–32. https://doi.org/10.11594/nstp.2022.2705

Urban, T. L. (2022). The effect of one-way aisles on retail layout. *Computers and Industrial Engineering*, *168*. https://doi.org/10.1016/j.cie.2022.108064

Valle, M. A., Ruz, G. A., & Morrás, R. (2018). Market basket analysis: Complementing association rules with minimum spanning trees. *Expert Systems with Applications*, *97*, 146–162. https://doi.org/10.1016/j.eswa.2017.12.028

Keywords: *Grocery Store Layout, Machine learning, Market Basket Analysis, Familly Business*