International Journal of Industrial Engineering & Production Research June 2024 Vol. 35, No. 2: 1-12
DOI: 10.22068/ijiepr.35.2.1967

RESEARCH PAPER

Multilevel Association Rules on Customers' Buying Pattern Based on Sales Transactions: A Case Study in Retail

Melinska Ayu Febrianti[1], Qurtubi[2*], Roaida Yanti[3] & Hari Purnomo[4]

Received 24 January 2024; Revised 30 March 2024; Accepted 24 April 2024;

ABSTRACT

The retail industry is a vital sector of the world economy and is characterized by fierce competition, tight profit margins, and demanding consumers. Understanding customer buying behavior patterns is essential in devising the best retail strategy to enhance product sales. This research aims to comprehend customer shopping behaviors based on retail sales transactions and formulate the best strategies. By employing multi-level association rules, the dataset is arranged hierarchically into categories, sub-categories, and items. The sales transaction data used comprises 5830 transaction records over a month. The results of this study reveal 24 associations of categories, 49 associations of sub-categories, and 12 associations of product items. Moreover, the proposed marketing strategy offers recommendations including store layout improvement, planogram design, and bundled product offerings. This research addresses the gap in empirical evidence from a previous study and suggests further observation from diverse locations to authenticate the findings, which may yield various outcomes.

KEYWORDS: *Buying behaviour; Data mining; FP-Growth; Marketing logistics; Retail strategies.*

1. Introduction

The retail industry is a vital economic sector worldwide, notified by more intense competition, tight profit margins, and demanding customers [1]. As a developing country, the acceleration of expansion in the Indonesian retail business is influenced by the people's buying power, growth, and need to fulfill consumption products [2]. The development of the retail industry in Indonesia is marked by the establishment of new outlets all over the region. The locations that are adjacent to each other reflect the rapid growth of the retail industry. It is highly related to people's shopping priorities: swiftness and practicality, especially in metropolitan areas where people have limited time and prefer to shop in a single location at a time.

In 2019, FMCG consumption in modern retail increased by 6.6%, whereas retail with the minimarket format experienced a growth of 12.1%. On the other hand, supermarkets and hypermarkets experienced a decrease of -6.8% [3]. Thus, the public prefers modern retail in the format of minimarkets to shop for daily necessities. With the growth of modern retail, retail company owners must strive to deal with the fiercer competition. Furthermore, the profit generated by this type of company is only about 7%-15% of sales [4]. Accordingly, the retail business must prioritize customer satisfaction as the most important factor in competitiveness.

Revealing customer's buying behavior is a primary focus in consumer behavior theory as an effort to prioritize consumer satisfaction. Numerous papers delve into consumer buying patterns on the internet. However, consumers' final purchasing decisions often change due to internal and external factors. Internal factors include cultural, social, personal, and psychological backgrounds. External factors include product, price, place, and promotion [5]. Therefore, it is essential for retailers, as businesses facing intense competition, to understand their customers' buying behavior.

ABC Retail is one of the retailers that sells everyday necessities and is located in the Sleman Regency, Special Region of Yogyakarta, Indonesia. This retail establishment was

* **Corresponding author: Qurtubi**
qurtubi@uii.ac.id

1. *Universitas Islam Indonesia*.
2. *Universitas Islam Indonesia*.
3. *Universitas Islam Indonesia*.
4. *Universitas Islam Indonesia*.

established ten years ago, and in February 2022, it opened its new branch. However, until now, the layout arrangement of the store and product allocation are not based on customer buying patterns but still rely on the owner's preferences. The poor arrangement of the store layout and shelf storage allocation can make it difficult for customers to find the items they want.

Optimizing the allocation of shelf space and store layout is required to maximize the visibility of products to customers and consequently stimulate impulsive purchases [1]. The location of the shelves on which the products are displayed significantly influences the sale of products. At the same time, displaying complementary products adjacent to each other can increase the possibility of cross-selling products (cross-selling) [6]. Therefore, the company's management must have and develop various management strategies and improve the suitable marketing method to survive in business and maintain the number of customers. Understanding customer buying patterns can help determine the best retail strategy to increase product sales.

Association Rule, or AR, is a well-known research tool for establishing product marketing strategies based on customer buying behavior. AR is finding patterns, correlations, relationships, or causal structures in a data set found in different types of databases, such as relational data, transactional data, and other types of data storage [7]. Market Basket Analysis (MBA) is an AR application often used to study customer purchasing habits. It is often called the Association Rule–Market Basket Analysis (AR-MBA). MBA is a technique used to collect observational data to understand consumer buying patterns within retail settings. Examining the contents of customer baskets and identifying connections between products aids retailers in shaping store layouts, devising diverse strategic initiatives, and determining merchandising strategies that have substantial effects on retail marketing and sales. According to Mostafa, an MBA analyzes customer buying behavior based on the correlations between several items purchased in a single transaction [8]. The MBA aims to inform the merchant about which goods the customer can buy simultaneously and which products are more suitable for marketing [9]. Since it can convey information from several levels of abstraction, multi-level can provide more specialized and focused information than single-level [10].

Rana and Mondal have conducted research related to consumer buying behavior with multilevel association-based approaches for market basket analysis and got information that provides a better understanding of consumer demand, which can be used as a basis for decision-making such as layout design, inventory control, merchandising activities, promotions, etc. In their research, Rana and Mondal used an apriori algorithm. They suggest future research to investigate faster alternative search techniques to reduce execution time and search space so that processing can be performed more quickly and accurately for a large data set [11].

Thus, the study discusses consumer behavior patterns to find multi-level association rules as one of the data mining methods by employing the Market Basket Analysis technique using the FP growth algorithm based on sales transaction data at the retail. In their research on market basket analysis using the apriori and FP growth algorithms, they found that the FP growth algorithm takes a shorter time than the apriori algorithm [12]. Furthermore, Chang et al. also compared FP growth with other algorithms. The results showed a significant reduction in the execution time of the item sets updated frequently in stages [13].

This research fills the empirical gap in the study of Randhawa and Saluja [14] and Pascucci et al. [15]. Randhawa and Saluja suggest future research to validate their findings by collecting data in different areas to produce varied results. Pascucci et al. recommend future research to repeat the study in other locations, types of store formats, and periods to verify the effectiveness of the proposed scorecard in category evaluating performance. Therefore, this research examines customer purchasing patterns in modern retail with different formats using the Multilevel Association Rule to validate product category performance and determine promotional policy strategies. This study also evaluates category performance as in the research by Pascucci et al. [15] with different locations, store formats, and periods. Still, this study does not utilize a scorecard but rather association rule mining.

2. Literature Review

2.1. Knowledge discovery in database (KDD)

According to Zanuardi and Suprayitno, KDD is a systematic analysis technique that finds patterns and obtains accurate, new, usable information from vast and complex data sets [16]. KDD has several steps: data selection, pre-processing, transformation, data mining, and evaluation/interpretation [17].

2.1.1. Data selection
Only data useful for research will be taken from the database for further processing, not from the entire database, which is often unemployed. The data are stored separately from the database.

2.1.2. Pre-processing
Data that was previously retrieved and selected is then cleaned, reduced, and integrated to check data discrepancies and eliminate incomplete data, data duplication, printing errors, and inconsistencies.

2.1.3. Transformation
Changing the selected data to be adjusted in a specific format following the software and algorithms used.

2.1.4. Data mining
The process of searching for interesting patterns or hidden information in the data that have been changed, based on the intended goal, with techniques such as classification, regression, clustering, or association rules and using the appropriate algorithmic rules. The entire KDD goal and process highly influence the right approach.

2.1.5. Evaluation/Interpretation
Re-examine whether the new information obtained contradicts or matches the previous fact or theory. Then, change the information from the data mining process into a method that interested parties can understand.

2.2. Association rule mining (ARM)
ARM, according to Triyanto is a data mining approach designed to find associative rules between items [18]. According to the research of Kim, Xu, and Gupta, association analysis became famous for its application in analyzing the buying habits of customers by searching for associations and correlations between different items in each customer's shopping cart, so association analysis is also known as market *basket* analysis [19].

Based on two measurements, support, and confidence, the association rule checks the number of times an item co-occurs in the transaction database [20]. Both measurements identify how often set items appear and their association rules. The rulemaking of association on the set item is if the support and confidence values are more significant than the minimum support and confidence values specified by the analyst [20]. Support indicates how often items appear in the database, while confidence means how often the statement is true [21]. Calculations are needed to evaluate the quality of association rules based on support, confidence, and lift ratio [22].

The multi-level association rule is a popular type of association rule. Multi-level association rules are employed to detect relationships of small groups of things in large volumes of data. Things are often classified hierarchically by category to deal with this issue. The association combinations of items are searched in stages using multi-level association rules in each hierarchical category. Thus, the interrelationships originating from the pair of items will be easier to detect.

2.3. FP-Growth algorithm
This algorithm is a modification of the Apriori algorithm in which the FP-Growth algorithm corrects deficiencies in the Apriori algorithm. The FP-Growth algorithm is more efficient and cost-effective because it saves time and storage space [20]. The FP-Growth algorithm does not generate candidate items to retrieve a combination of items from the database, which is the difference between these two techniques [21]. Generating candidates ensures that the candidate set is included in the frequent itemsets. Meanwhile, the FP-Growth algorithm uses the FP-Tree development idea to search for routine item sets. The FP-Growth algorithm searches for the most common item sets in the studied data.

Tab. 1. literature review

No.	Previous research	The differences between this research
1	Rana and Mondal [11]	The research does not provide a marketing strategy based on the results produced.
2	Nurmayanti et al. [23]	The research only investigates single level associations, not multilevel
3	Firmansyah [24]	The research was conducted in bookstores and only examined single levels of association rules.
4	Sağın and Ayvaz [25]	The research investigates single level associations and comparing the results of two methods between the apriori and fp-growth algorithms.
5	Raorane et al. [26]	The research investigates single level and does not provide a

No.	Previous research	The differences between this research
		marketing strategy based on the results produced
6	Malik et al. [27]	The research investigates single level and used machine learning.
7	Singha et al. [28]	The research was conducted in health food store and the marketing strategy provided focuses on cross selling and product discounts.

3. Research Method

3.1. Research methods

The object of this study is data on sales transactions at ABC retail, which sells a wide range of daily necessities products. This study analyzes customer shopping behavior patterns in retail using multilevel association rule mining with the FP-Growth algorithm. Multilevel association rule mining is a variation of association rule mining that discovers relationships among items at each level by applying different thresholds at different levels [29]. Therefore, searching for associations from item combinations is carried out gradually at each hierarchical category level, making detecting relationships among item pairs easier. The information provided by multilevel mining will be more significant, specific, and focused compared to the single-level mining [30].

In this study, the sales transaction data used is one month. The research variables include the transaction code and product name, while other variables are only used as additional information. The criteria for a sales transaction is having more than 1 item in each transaction. The total transaction data obtained amounted to 5,830 transactions.

Then, the data is processed based on the KDD process. The first step is selection, selecting the data obtained. The second is data pre-processing, which involves several data preparation processes before the data mining. The third is Transformation, where consumer transaction data is transformed into data per transaction in each row. So that the data can be processed further using RapidMiner software.

Fourth is data mining, which is data processing based on FP growth algorithms according to data mining techniques using RapidMiner software. The procedure for determining frequent itemsets is divided into two steps [21], constructing the FP-Tree, and applying the FP-Growth algorithm. Constructing the FP-Tree involves mapping transaction data onto specific paths in the FP-Tree because there may be transactions with the same items in each transaction; thus, compressing paths with the FP-Tree data structure is more effective. Then, applying the FP-Growth Algorithm begins with generating a conditional pattern base, a sub-database containing prefix and suffix patterns for paths. Subsequently, conditional FP-tree generation is carried out on the number of supports more significant than the minimum support count using the FP-tree. Finally, the frequent item sets are searched by combining derived items for each conditional FP-tree.

The last step is interpretation evaluation, which interprets the association results obtained from data mining techniques. Furthermore, the evaluation process is seen from the algorithm parameters (support, confidence, lift). So that new information is obtained based on the results of the analysis.

3.2. Product categorization

The in-store products are divided into more specific categories. The categorization of products is determined based on the similarity between the products. In this study, the categorization of products refers to research conducted by Lusiani [4]. Of many products sold, it will be divided into 28 categories and 127 sub-categories. Details of the categories and categories can be found in Table II.

Tab. 2. list of product categories and sub-categories

No.	Category	Sub-categories
1	Beauty Kit	Comb; Nail clippers; Accessories; Pingset
2	Apparel	Underwear; Socks; Towel
3	Baby	Diapers; Milk bottles; Baby food
4	Breakfast	Ground coffee; Tea; Cereals; Milk powder; Powdered drinks; SCM
5	Cigarette	Cigarette
6	Cleaner	Detergents; Dish soap; Floor cleaner; Deodorizer
7	Condiment	Salt; Msg; Soy sauce; Sauces; Vinegar; Mayonnaise; Coconut milk; Food Coloring; Vanilla; Jelly powder
8	Confectionery	Sweets; Chocolate

No.	Category	Sub-categories
9	Cooking Oil	Oil; Butter; Margarine
10	Dairy	UHT milk; Cheese; Ice cream; Yogurt
11	Egg	Egg
12	Electrical	Batteries; Lamps; Cable
13	Entertainment	Toys; Sports; Birthday fixtures
14	Fruit	pears; Banana
15	Healthcare	Medications; Vitamins; Plaster; Condoms; Patch
16	Houseware	Cutlery; Mats; Napkins; Broom; Mop; Oil paper; Washing brush; Sponge
17	Insecticide & Air freshener	Insect repellent; Seodorizers
18	Instant Food	Instant noodles; Canned food; Frozen food; Food sachets
19	Jam & Spread	Jam; Sprinkles; Honey
20	Miscellaneous	Matches; Pulses; Umbrella; Raincoats; Slippers; Scissors; Plastic; Gallons of water; Rope; Rubber; Polish; Candles; Straw
21	Drink	Soft drink; Packaged tea; Beer; Juice; Jelly drinks; Mineral water; Health drinks; Liquid coffee
22	Ready to eat	Bread; Cakes; Sausage
23	Rice-flour-sugar	Rice; Flour; Sugar; Sago
24	Sanitary	Dressings; Wipes; Toothpicks; Cotton; Cotton buds
25	Snack	Modern; Traditional; Biscuits; Wafers; Nuts; Cryptic; Krupuk; Jelly
26	Stationery	Stationery; Wrapping paper; Books; Glue
27	Toiletries	Haircare; Oral care; Skin care; Deodorant; Shavers; Perfumes; Body care
28	Gasoline-water gallons-gas	Gasoline; Water Gallons; Gas

4. Data Processing

4.1. Pre-processing category data

In this study, three stages were carried out in the pre-processing data. The first stage is data cleaning by removing failed transaction data that exclude essential components, such as item names, dates, and item codes, as well as transactions that only contain 1 type of product. After data cleaning is performed towards 5830 transaction data, 2602 data are obtained for processing by category. Furthermore, the data reduction stage reduces several insignificant variables in the transaction data, including the variable date, item code, unit price, and the number of product items sold. So, the variables used are only the transaction number and the name of the product item. Then, the third stage is to integrate data by combining products of the same type to become a unified product group. Table III demonstrates the results of the pre-processing data stage by category.

Tab. 3. Pre-processing Category Data

Transaction	Category
1	Breakfast; Cleaner; Condiment; Snack; Toiletries
2	Condiment; Dairy
3	Condiment; Instant food
4	Ready to eat; Snack
5	Cigarettes; Drink; Snack
...	...
2801	Drink; Ready to eat

4.2. Category data transformation

Data should be transformed first before carrying out the data mining process to find associations. Data transformation is carried out to adjust the data to the software used for processing. In this study, data processing will use RapidMiner software to transform transaction data into binary numbers, namely 0 and 1.

4.3. Category data parameter determination

According to Larose, researchers can determine the minimum value of support and minimum confidence as needed [31]. This study uses the trial and error method for parameter determination based on the associations formed with the FP-Growth algorithm. Table IV shows the results of experiments to determine this study's minimum value of support and confidence.

Tab. 4. Determination experiments for support and confidence parameter

Experiment	Min. support	Min. Confidence	Result
1	0.1	0.5	No rule is formed
2	0.01	0.5	There are rules with slight variations
3	0.02	0.2	There are rules with many variations

Based on the experiments carried out in this research, the selected values of the support and confidence parameters are by the 3rd experiment. The support value is recorded as 0.02 or 2%, and the confidence value is determined as 0.2 or 20%. The value will then be used to determine the association using the FP-Growth algorithm.

5. Result and Discussion

Product buying Patterns in this research were analyzed quantitatively using data mining methods on Customer Baskets. The results show rules that provide an overview of the patterns that emerge, which particular items are more related to each other, and the relationships between them and their importance.

5.1. Identification of associations between categories

Table V presents the results of associations formed between categories using RapidMiner software.

Tab. 5. Association results between categories

No	Premises	Conclusion	Support	Confidence	Lift
1	Instant Food	Ready to eat	0,03	0,20	1,068
2	Confectionery	Dairy	0,03	0,22	1,462
3	Breakfast	Instant food	0,04	0,23	1,516
4	Snack	Ready to eat	0,10	0,23	1,222
5	Confectionery	Ready to eat	0,03	0,23	1,230
6	Toiletries	Cleaner	0,03	0,24	2,618
7	Cleaner	Breakfast	0,02	0,24	1,366
8	Dairy	Ready to eat	0,04	0,25	1,329
9	Condiment	Breakfast	0,02	0,26	1,456
10	Instant Food	Breakfast	0,04	0,27	1,516
11	Condiment	Instant food	0,02	0,30	1,979
12	Sanitary	Toiletries	0,02	0,30	2,097
13	Cleaner	Snack	0,03	0,34	1,818
14	Drink	Cigarette	0,12	0,35	1,097
15	Rice-flour-sugar	Breakfast	0,02	0,38	2,130
16	Cleaner	Toiletries	0,03	0,38	2,618
17	Miscellaneous	Cigarette	0,02	0,38	1,208
18	Cigarette	Drink	0,12	0,39	1,097
19	Gasoline-Gallons-Gas	Cigarette	0,03	0,43	1,356
20	Instant Food	Snack	0,07	0,48	1,159
21	Drink, Ready to eat	Snack	0,03	0,50	1,196
22	Ready to eat	Snack	0,10	0,51	1,222
23	Dairy	Snack	0,08	0,52	1,233
24	Confectionery	Snack	0,07	0,55	1,304

From the data processing results in Table V, 24 associations between categories were obtained from the minimum value limit of support of 2% and the minimum confidence of 20%. The support value indicates that the frequent category pairs appear in at least 2% of transactions. Out of the 2602 transactions, the minimum appearance of a category pair is "frequent" if recorded in as many as 52 transactions. The small value of support aims to identify more associations for analysis purposes. The confidence value indicates that the confidence level of the predecessor category and the follower category will appear on the same transaction at least 20%.

Moreover, from the 24 associations formed, the best association rule based on the highest lift value is the association between the cleaner and toiletries categories. This association has a lift value of 2.618 with a confidence value of 38% and a support value of 3%. Meanwhile, the association rule with the smallest lift value is the association between the instant food and ready-to-eat categories, which has a lift value of 1.068 with a confidence of 20% and a support of 3%. Valle, Ruz, and Morras [32] stated in their research that researchers can rank association rules from the best to the weakest based on the highest to lowest lift values.

5.2. Identification of associations between sub categories

In data processing, the previous stage's results, namely category associations, are used to find sub-category associations. At this stage, only sub-categories data belonging to those associated categories are employed. Table VI presents the results of associations formed between sub-categories.

Tab. 6. Results of association between sub-categories

Category Association to-	Premises	Conclusion	Support	Confidence	Lift
1	Instant noodles	Cake	0,205	0,231	1,128
2	Candy	Ice cream	0,230	0,354	1,011
3	Powdered drinks	Instant noodles	0,245	0,963	1,001
	Ground coffee	Instant noodles	0,547	0,967	1,005
4	Wafer	Cake	0,100	0,296	1,128
	Biscuit	Cake	0,138	0,319	1,214
	Biscuit	Sausage	0,169	0,391	1,079
	Wafer	Sausage	0,138	0,407	1,123
	Traditional	Bread	0,119	0,500	1,159
5	Chocolate	Cake	0,229	0,647	1,826
6	Oral care	Dish soap	0,101	0,333	1,648
7	Detergent	Powdered drinks	0,263	0,341	1,079
	Detergent	Ground coffee	0,368	0,477	1,088
8	UHT milk	Sausage	0,259	0,344	1,115
9	Sauce	Ground coffee	0,130	0,412	1,307
	Msg	Powdered drinks	0,185	0,417	1,607
	Soy sauce	These	0,111	0,750	2,892
10	Instant noodles	Powdered drinks	0,245	0,255	1,000
	Instant noodles	Ground coffee	0,547	0,569	1,004
11	Sauce	Instant noodles	0,654	0,919	1,016
12	Tissue	Body care	0,264	0,560	1,187
13	Detergent	Traditional	0,250	0,370	1,234
14	Mineral water	Cigarette	0,285	1,000	1,000
	Packaged tea	Cigarette	0,235	1,000	1,000
	Healthy drinks	Cigarette	0,223	1,000	1,000
	Liquid coffee	Cigarette	0,210	1,000	1,000
15	Sugar	Tea	0,283	0,378	1,133

Category Association to-	Premises	Conclusion	Support	Confidence	Lift
16	Sugar	Ground coffee	0,450	0,600	10,280
	Oral care	Dish soap	0,101	0,333	1,648
17	Match	Cigarette	0,758	1,000	1,000
18	Cigarette	Liquid coffee	0,210	0,210	1,000
	Cigarette	Healthy drinks	0,223	0,223	1,000
	Cigarette	Packaged tea	0,235	0,235	1,000
	Cigarette	Mineral water	0,285	0,285	1,000
19	Gasoline	Cigarette	0,945	1,000	1,000
20	Instant Noodles	Traditional	0,258	0,271	1,008
	Instant Noodles	Wafer	0,268	0,282	1,010
	Instant Noodles	Biscuit	0,284	0,298	1,012
21	bread, mineral water	Biscuit	0,167	0,563	1,215
22	Bread	Traditional	0,119	0,275	1,159
	Sausage	Wafer	0,138	0,379	1,123
	Cake	Wafer	0,100	0,381	1,128
	Sausage	Chips	0,144	0,397	1,442
	Sausage	Biscuit	0,169	0,466	1,079
	Cake	Biscuit	0,138	0,524	1,214
23	UHT milk	Wafer	0,205	0,322	1,098
	UHT milk	Biscuit	0,258	0,406	1,015
24	Candy	Chips	0,249	0,338	1,055
	Candy	Wafer	0,287	0,391	1,056

Table VI shows the data processing results for sub-categories obtained from 24 category associations in the previous process, with 49 associations between sub-categories. These sub-category association rules can be used to make the product layout in a department more specific and focused. Furthermore, Table VII presents the associations of the sub-categories with the highest support values.

Tab. 7. Association of Sub Categories with the Highest Support Value

No	Premises	Conclusion	Support	Confidence	Lift
1	Gasoline	Cigarette	0,945	1,000	1,000
2	Match	Cigarette	0,758	1,000	1,000
3	Sauce	Instant noodles	0,654	0,919	1,016
4	Ground coffee	Instant noodles	0,547	0,967	1,005

Based on Table VII, four associations were obtained to identify product item associations: gasoline with cigarettes matches with cigarettes, sauces with instant noodles, and ground coffee with instant noodles.

5.3. Identification of associations between product items

In data processing, the previous stage results are used to find product item associations, namely sub-category associations. Product items are the lowest level of the data set hierarchy when identifying multilevel association rules. Only the product items included in the four sub-category associations with the highest support values are employed at this stage. Table VIII presents the results of the associations formed between product items.

Tab. 8. Results of Association between Product Items

Association of Sub-Categories	Premises	Conclusion	Support	Confidence	Lift
1	Mild 16	Gasoline pertalite	0,087	1,000	1,000
	Cigarette retail 1 pcs	Gasoline pertalite	0,072	1,000	1,000
	Kretek orbit 12	Gasoline pertalite	0,058	1,000	1,000
	Gg solar 12	Gasoline pertalite	0,058	1,000	1,000
	Aspro international 16	Gasoline pertalite	0,058	1,000	1,000
2	Diplomat evo 16	Tokai match	0,106	1,000	1,306
	Gg solar 12	Tokai match	0,106	0,833	1,088
3	Sauce in sct /3	Fry 75g	0,059	0,500	8,500
	ABC sambal ex pds 80g	Instant noodled soto 70g	0,059	1,000	6,800
4	Coffemix 20g	Instant noodles in cup - chicken onion 28g	0,055	0,429	7,857
	Coffemix 20g	Oxtail soup noodles in cup	0,055	0,429	7,857
	Good day chocochino 20g	Instant noodles grg sp 85g	0,091	0,833	2,865

A retailer can use this product item association to increase product sales and, in turn, its revenues by using these rules in catalog design, in-store layout, warehouse design, discount promotion arrangements, and decision support system design [25].

5.4. Marketing strategy recommendations

The pattern of buying behavior yielded from RapidMiner output states which products are most often purchased simultaneously at ABC retail. Based on the results of the association, proposal of marketing strategies are recommended to increase sales as follows:

5.4.1. Arrangement in-store layout

Improving the store layout aims to simplify activities for buyers and sellers to achieve efficiency and effectiveness. Retailers and stores are growing everywhere, so customers expect to experience good service from retailers. The faster the retailer provides service, the more satisfied the customer will be. In addition, the store's layout arrangement can create a track across the entire physical location that can guide consumers through merchandise products they want to sell more.

5.4.2. The allocation management for display rack space planogram

The implementation of display shelf space management is necessary to optimize the allocation of product diversity on limited display shelf space. Space management, commonly called a planogram, is a concept or plan to display products based on consumer spending habits to maximize profits and improve consumer services. Product placement is essential for a retail business owner in allocating products on the shelves since each product has a difference in sales if the product is distributed at a different level.

5.4.3. Bundling product offerings

Associated products can be offered to consumers by selling these products in one bundle. Cross-selling product offerings with products sold in a single bag will be more attractive if schemes benefit consumers, such as providing discounts or bonuses for other products. In determining the scheme, it is necessary to conduct further research that considers the price factor in determining product associations. According to Aazami and Saidi-Mehrabad [33] price is a crucial factor when developing the demand function. Based on the analysis, some products offered are sauces with instant noodles, ground coffee and instant noodles, sugar and ground coffee, detergents and ground coffee, sweets, wafers, mineral water, and cigarettes.

The managerial insight.

This paper offers several implications for retail managers, particularly those in the physical store sector. This research validates Randhawa and Saluja's research that merchandising strategy can influence customer behavior. Of the six merchandising factors analyzed by Randhawa and Saluja, there are three factors related to this research: shelf presentation and product display, store layout, and promotional schemes and discount offers.

The proposed market basket analysis introduces measures of association rules from a theoretical standpoint, which capture the monetary significance of associations and the frequency of item set occurrences. These metrics aid users in evaluating marketing strategies by assessing their impact on overall company revenue through the associations they generate. Moreover, crafting effective marketing strategies involves considering customer purchasing patterns and allocating the company's marketing budget. Strategies suggested for implementation, based on frequent and high-value itemset associations, encompass adjustments to store layout, management of planogram allocation, and bundling products to facilitate cross-selling. Professionals can utilize the broad analytical framework to apply the proposed market basket analysis to diverse transactional datasets across various industries.

6. Conclusion

This article identifies purchasing behavior patterns through basket analysis employing multilevel and FP-Growth algorithms on transaction records of products from one of the retailers. The study uncovers common associations in customer purchasing behavior to design appropriate marketing strategies, such as layout management, planograms, and cross-selling. In addition to measuring the commonly used performance of the generated associations, our research also suggests the importance of marketing strategies that focus on discounts and promotions, in-store layout arrangements, and allocation management for display rack space planograms. Further research is expected to analyze customer behavior patterns with a more significant amount of transaction data and a more extended period with different research subjects. Future research can analyze customer satisfaction and increase sales after changing the retail layout and planogram design based on association rules. In addition, further research can design store layouts and planograms based on the analysis of association rules made in this study.

References

[1] T. Flamand, A. Ghoniem and B. Maddah, "Promoting Impulse Buying by Allocating Retail Shelf Space to Grouped Product Categories," *Journal of the Operational Research Society*, (2016), pp. 953-969.

[2] R. Y. Eldianty, "Pengaruh Perceived Quality, Commercial Image, Price Image Terhadap Purchase Intention Melalui Brand Equity Produk Store Brand di Hypermart dan Supermarket di Surabaya," *Calyptra: Jurnal Ilmiah Mahasiswa Universitas Surabaya*, Vol. 7, No. 1, (2018), pp. 1788-1802.

[3] R. Hadyan, "Kemendag: Industri Ritel Modern Masih Menunjukkan Tren Positif," 27 Agustus 2019. [Online]. Available: https://www.kemendag.go.id/id/newsroom/media-corner/kemendag-industri-ritel-modern-masih-menunjukkan-tren-positif-1. Accessed 3 Februari (2022).

[4] M. Lusiani, "Optimasi Alokasi Produk pada Ruang Rak Displai Gerai Minimarket berdasarkan Harga Produk Menggunakan Multilevel Association Rules," *JIEMS (Journal of Industrial Engineering and Management Systems)*, Vol. 4, No. 2, (2017), pp. 15-32.

[5] H.Chang, L. Huang and C. Ho., "An anticipation model of potential customers' purchasing behavior based on clustering analysis and association rules analysis," *Expert Systems with Applications* , Vol. 32, (2007), p. 753-764.

[6] T. Ozcan and S. Esnaf, "A discrete constrained optimization using genetic algorithms for a bookstore layou," *International Journal of Computational Intelligence Systems*, Vol. 6, No. 2, (2013), pp. 261-278.

[7] P. Dhanalakshmi and R. Porkodi, "A Survey on Different Association Rule Mining Algorithms in Data Mining," *IPASJ International Journal of Computer Science,* (2017), pp. 126-133.

[8] M. M. Mostafa, "Knowledge discovery of hidden consumer purchase behaviour: a market basket analysis," *International Journal of Data Analysis Techniques and Strategies*, Vol. 7, No. 4, (2015), pp. 384-405.

[9] B. Umayah and F. Kurniawan, "Analisa Perilaku Konsumen Melalui Data Transaksi Berbasis Pendekatan Market Basket Analysis," *Sains, Aplikasi,*

Komputasi dan Teknologi Informasi, Vol. 1, No. 2, (2019), pp. 30-35.

[10] T. Pradana, "Penggalian Kaidah Multilevel Association Rule dari Data Mart Swalayan ASGAP," *Spirit: STMIK Yadika Journal of Computing and Cybernetic System,* Vol. 7, No. 2, (2015).

[11] S. Rana and M. N. I. Mondal, "“A Seasonal and Multilevel Association Based Approach for Market Basket Analysis in Retail Supermarket,”," *European Journal ofInformation Technologies and Computer Science,* Vols. Vol. 1, No. 4, (2021).

[12] M. Hossain, A. H. M. S. Sattar and M. K. Paul, "“Market Basket Analysis Using Apriori and FP Growth Algorithm,”," in *2019 22nd International Conference on Computer and Information Technology (ICCIT),* (2019).

[13] H. Y. Chang, J. C. Lin, M. L. Cheng and S. C. Huang, "“A novel incremental data mining algorithm based on FP-growth for Big Data,”," in *2016 International Conference on Networking and Network Applications (NaNA),* (2016), pp. 375-378.

[14] K. Randhawa and R. Saluja, "Does Visual Merchandising have an Effect on Consumer Impulse Buying Behavior? A Study with Special Reference to Apparels in Punjab," *Journal of General Management Research,* Vol. 4, No. 2, (2017), pp. 58-71.

[15] F. Pascucci, L. Nardi, L. Marinelli, M. Paolanti, E. Frontoni and G. Gregori, "Combining sell-out data with shopper behaviour data for category performance measurement: The role of category conversion power," *Journal of Retailing and Consumer Services,* Vol. 65, (2022).

[16] A. Zanuardi and H. Suprayitno, "Analisa karakteristik kecelakaan lalu lintas di jalan ahmad yani surabaya melalui pendekatan knowledge discovery in database," *Jurnal Manajemen Aset Infrastruktur & Fasilitas,* Vol. 2, No. 1, (2018).

[17] R. R. Putra and C. Wadisman, "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means," *INTECOMS: Journal of Information Technology and Computer Science,* Vol. 1, No. 1, (2018), pp. 72-77.

[18] W. A. Triyanto, "Association Rule Mining Untuk Penentuan Rekomendasi Promosi Produk," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer,* Vol. 5, No. 2, (2014), pp. 121-126.

[19] H. W. Kim, Y. Xu and S. Gupta, "Which is more important in Internet shopping, perceived price or trust?," *Electronic commerce research and applications,* Vol. 11, No. 3, (2012), pp. 241-252.

[20] M. Nafari and J. Shahrabi, "A temporal data mining approach for shelf-space allocation with consideration of product price," *Expert Systems with Applications,* Vol. 37, No. 6, (2010), pp. 4066-4072.

[21] G. Gunadi and D. I. Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Penjualan Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (FP-Growth)," *Telematika Mkom,* Vol. 4, No. 1, (2012).

[22] D. Firdhana, Penerapan Metode Market Basket Analysis untuk Sistem Costumer, STIKOM, (2006).

[23] W. P. Nurmayanti, H. M. Sastriana, A. Rahim, M. Gazali, R. H. Hirzi, Z. Ramdani and M. Malthuf, "Market basket analysis with apriori algorithm and frequent pattern growth (Fp-Growth) on outdoor product sales data.," *International Journal of Educational Research & Social Sciences,* Vol. 2, No. 1, (2021), pp. 132-139.

[24] F. Firmansyah, "Market Basket Analysis for Books Sales Promotion using FP Growth Algorithm, Case Study: Gramedia

Matraman Jakarta," *Journal of Informatics and Telecommunication Engineering,* Vol. 4, No. 2, (2021), pp. 383-392.

[25] A. N. Sagin and B. Ayvaz, "Determination of association rules with market basket analysis: application in the retail sector," *Southeast Europe Journal of Soft Computing,* Vol. 7, No. 1, (2018).

[26] A. A. Raorane, R. V. Kulkarni and B. D. Jitkar, "Association rule–extracting knowledge using market basket analysis," *Research Journal of Recent Sciences,* Vol. 1, No. 2, (2012), pp. 19-27.

[27] M. H. Malik, H. Ghous, M. Ismail, S. Jamshaid and J. Altaf., "Market Basket Analysis for Next Basket Item Prediction Using Data Mining and Machine Learning," *Journal of Computing & Biomedical Informatics* , (2024).

[28] K. Singha, P. Parthanadee, A. Kessuvan and J. Buddhakulsomsiri, "Market Basket Analysis of a Health Food Store in Thailand: A Case Study," *International Journal of Knowledge and Systems Science,* Vol. 15, No. 1, (2024), pp. 1-14.

[29] R. Agarwal., "nventory Classification Using Multi-Level Association Rule Mining," *International Journal of Decision Support System Technology,* Vol. 11, No. 2, (2019), pp. 1-12.

[30] T. Pradana, "Penggalian Kaidah Multilevel Association Rule dari Data Mart Swalayan ASGAP," *Spirit: STMIK Yadika Journal of Computing and Cybernetic System,* Vol. 7, No. 2, (2015).

[31] D. T. Larose and C. D. Larose, Discovering knowledge in data: an introduction to data mining, New Jersey: John Wiley & Sons, (2014).

[32] M. Valle, G. Ruz and R. Morras., "Market Basket Analysis: Complementing Association Rules With Minimum Spanning Trees," *Expert System With Application,* Vol. 97, (2018), pp. 146-162.

[33] A. Aazami and M. Saidi-Mehrabad., "A production and distribution planning of perishable products with a fixed lifetime under vertical competition in the seller-buyer systems: A real-world application," *Journal of manufacturing systems,* Vol. 58, (2021), pp. 223-247.

Follow this article at the following site:

Melinska Ayu Febrianti, Qurtubi, Roaida Yanti & Hari Purnomo. Multilevel Association Rules on Customers' Buying Pattern Based on Sales Transactions: A Study Case in Retail. IJIEPR 2024; 35 (2) :1-12
URL: http://ijiepr.iust.ac.ir/article-1-1967-en.html