Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

Market Basket Analysis to Determine Muslim Clothing Supply in Indonesia Ahead of Eid Al-Fitr

Gun Gun Indra Gunawan*, Tri Wahyu Aji, Christina Juliane

School of Business and Information Technology, STMIK LIKMI, Bandung, Indonesia
Email: [1]gun.ig97@gmail.com, [2]triwaji@gmail.com, [3]christina.juliane@likmi.ac.id
Correspondence Author Email: gun.ig97@gmail.com
Submitted: **10/05/2024**; Accepted: **23/06/2024**; Published**: 23/06/2024**

Abstract−Enterprise transaction data is a valuable source of insights for companies to increase sales. In preparation for Eid al-Fitr, this study leverages Market Basket Analysis with the FP-Growth algorithm to uncover buying patterns within Indonesia's Muslim clothing market. Market Basket Analysis is one way to explore information through data to find customer buying patterns that are often used as insight into company decision-making. The data processing method uses the FP-Growth algorithm, which generates association rules based on calculating the frequency of occurrence of itemsets. Using the FP-Growth algorithm gives good results in the determination of association rules. From Muslim fashion store transaction data over the last 12 months, it produced 30 item set patterns with a minimum support value of 0.009 and confidence of 0.58. By identifying these in-demand product pairings, businesses can make informed decisions about stock allocation. This ensures they have the right combination of items available to meet customer needs during the surge in demand leading up to Eid al-Fitr. Additionally, these patterns can inform targeted promotional campaigns and strategic bundling initiatives, maximizing sales and customer satisfaction throughout this critical sales period.

Keywords: FP-Growth; Association Rules; Market Basket Analysis; Muslim Clothing Stores; Eid al-Fitr

1. INTRODUCTION

Thousands or even millions of transaction data are recorded daily in a company, including data from various fields such as marketing, finance, health, etc. Currently, the use of manual recording to store all transaction information has decreased due to technological advances. Companies record and store all their transaction history in a database. As time goes by, the amount of data stored will increase. This data is an essential asset for the company. Companies can make crucial decisions through data, such as strategies to increase sales, manage stock, and maintain customer loyalty [1]. Since this database consists of tens of thousands of records, it is impossible to read it manually and look up information from it. Therefore, various data mining techniques (association, clustering, classification, and prediction) are used [2].

Market Basket Analysis is one of the data mining techniques used to find customer buying patterns in a company [3]. The goal is to recommend to customers products to buy, promote, and increase sales. The pattern is observed when considering products that customers often purchase together. In the process of market basket analysis, methods or algorithms are needed to help the analysis process. In previous research, A. Setiawan and R. Mulyanti. (2020) [4] used the Apriori algorithm to analyze shopping baskets in trendy Muslim fashion stores. Currently, the most widely used methods for Market Basket Analysis are FP-Growth and Apriori algorithms [5], [6], [7]. Noviana et al. (2023) [8] compare Apriori and FP-Growth algorithms in determining consumer buying patterns. Then Aldino et al. (2023) [9] also conducts a comparative market basket analysis to determine consumer buying patterns using FP-Growth and Apriori algorithms. Based on the results of both studies, the FP-Growth algorithm is superior in computational time to the a priori algorithm Given this significant advantage, this study leverages the FP-Growth algorithm to conduct Market Basket Analysis (MBA) on customer transaction data. MBA is a powerful technique for uncovering hidden patterns in customer buying behavior. By analyzing frequent itemsets – combinations of products that customers frequently purchase together – MBA can reveal valuable insights into customer preferences, product associations, and potential upselling or cross-selling opportunities. This study employs the FP-Growth algorithm to achieve a more efficient and scalable analysis of customer buying patterns. This allows for a deeper understanding of customer behavior and ultimately leads to more informed and strategic business decisions. The insights gleaned from MBA can be used to optimize product placement, develop targeted marketing campaigns, and ultimately increase customer satisfaction and revenue.

The FP-Growth algorithm is more efficient, scalable, and memory-efficient and generates more accurate association rules than Apriori algorithms [10]. This algorithm can determine the items that appear most frequently in a dataset and determine association rules (relationships between those items) [11]. The collection of items will be determined by creating a data structure in a tree called FP-Tree. FP-Growth avoids "candidate generation" by building a data structure called FP-Tree [12]. FP-Tree is used to calculate frequent itemset support directly, making it more efficient in more significant amounts of data. This method has also been used in research by Saputra et al. (2023) [13] in conducting Market Basket Analysis to determine customer buying patterns.

This study ventures into the dynamic realm of the Muslim fashion industry, aiming to bridge the gap between retail inventory management and a nuanced understanding of consumer behavior. Employing market basket analysis, specifically the FP-Growth algorithm, delves into the fascinating world of customer buying patterns within Muslim fashion stores. This research transcends a simple analysis of transaction data. It incorporates valuable customer

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

information, weaving a richer tapestry that reveals the intricate relationships between Muslim clothing items frequently purchased together. The cornerstone of this study lies in the meticulous cleaning and processing of data. This ensures the analysis rests upon a foundation of accuracy, allowing for extracting the most valuable insights. The research unveils the hidden language of consumer purchase preferences by leveraging the power of association rules. It sheds light on the fascinating patterns of co-occurrence, illuminating which Muslim clothing items are most likely to be purchased together. This knowledge empowers Muslim fashion store owners with a strategic advantage. They gain invaluable insights that inform critical decisions across various aspects of their business. Inventory management becomes a more precise science, with stock levels optimized to meet the demands revealed by customer buying patterns. Product placement strategies can be tailored to leverage the power of suggestion, strategically positioning complementary items to entice additional purchases. Targeted marketing campaigns become laser-focused, reaching the right customer with the right product at the opportune moment. Ultimately, the findings of this study empower Muslim fashion businesses to cater more effectively to their customer's evolving needs and desires. This translates to increased sales and a surge in customer satisfaction, particularly during peak seasons like Eid al-Fitr. By bridging the gap between data and customer behavior, this research offers a roadmap to success, enabling Muslim fashion businesses to flourish in a competitive and ever-changing landscape. The next part of the article is structured: the study and methodology are presented in Section. 2, results and discussion are discussed in Section. 3, and the conclusions are detailed in Section. 4.

2. RESEARCH METHODOLOGY

The research methodology summarizes the steps to conduct research using the FP-Growth algorithm in Market Basket Analysis to determine customer buying patterns. The stages of this study consist of data collection, preprocessing, data processing using the FP-Growth algorithm, and then analysis and conclusions are drawn on the research results. A flowchart of research steps can be seen in Figure 1.

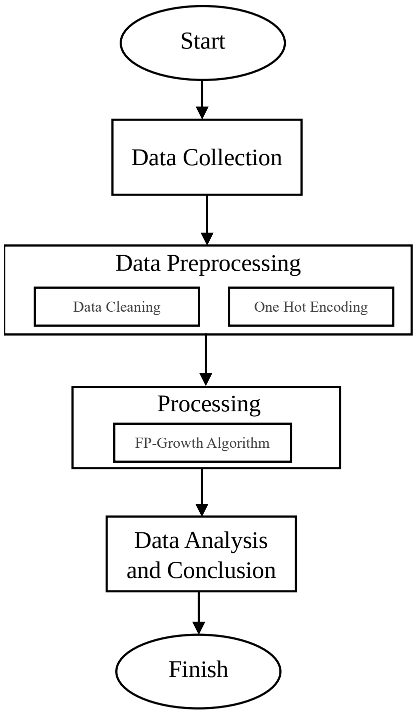

Figure 1. Flowchart market basket analysis with fp-growth algorithm

2.1 Data Collection

To understand customer buying behavior comprehensively, this study utilizes a rich dataset encompassing 12 months of transaction data collected at Muslim clothing store X. The data covers an entire year, from January 1, 2023, to December 31, 2023, capturing seasonal trends and variations in customer preferences. This extended timeframe allows for more robust analysis and the identification of potential patterns specific to different seasons, such as the surge in demand for traditional clothing items around Eid al-Fitr.

2.1.1 Data Description

This study leverages a substantial dataset encompassing 3,085 orders for Muslim clothing collected from store X. This extensive data collection effort provides a robust foundation for analyzing customer buying patterns. The dataset spans an entire year, offering valuable insights into seasonal trends and potential variations in customer preferences. Each

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

order within the dataset is uniquely identified by an "OrderID," allowing for precise tracking and analysis of individual transactions. The dataset encompasses a comprehensive range of information across 34 columns, including details on purchase transactions and customer profiles. While various data points are captured, the market basket analysis focuses on two key columns: "OrderID" and "Product Details." "OrderID" serves as the unique identifier for each transaction, enabling the differentiation and analysis of individual purchases. "Product Details" provides crucial information regarding each order's specific Muslim clothing items. This detailed product information forms the core element for frequently identifying co-purchased items and uncovering association rules. At the same time, Product Details are details of products purchased in each order. A preview of the data can be seen in Table 1.

Table 1. Sales dataset

No	Order ID	Product Details
0	#19141	Zayyan Dewasa Black L Black x 1 \n Zayyan Dewasa Brown Olive L Brown Olive x 1
1	#19140	Latif 07 Pendek Abu L Abu x 1
2	#19139	Rairaka 06 Denim 10 Denim x 1 \n Rairaka 06 Matcha 6 Matcha x 1
3	#19138	Kurta Ahnaf Anak Dark Grey XXS Dark Grey x 1
4	#19137	Kurta Ahnaf Dewasa Black S Black x 1 \n Kurta Ahnaf Dewasa Black L Black x 2
...	...	...
3084	#15955	Shaka Dewasa Black M Black x 5\n Shaka Dewasa ...

2.2 Data Preprocessing

2.2.1 Data Cleaning

The data cleaning is crucial to ensure the accuracy and effectiveness of the market basket analysis. This process involves three distinct stages. The first stage tackles the challenge of separating multiple purchased items listed together in the "Product Details" column. Here, we meticulously split each row in this column into separate entries for each item. This separation utilizes a newline character ("\n") as a delimiter, creating a new row for each item within a transaction. The Python function split() proves instrumental in achieving this separation, transforming the combined product details into a list containing each item as a distinct element. Following the separation stage, we address the potential issue of leading or trailing spaces within individual items. These unnecessary whitespace characters could hinder analysis. This stage meticulously removes any such spaces from the beginning and end of each item within each transaction. The Python function strip() is invaluable for this task, ensuring consistency and eliminating extraneous spaces for each item in every list. To achieve this across all transactions, the apply() function is employed, iteratively applying the strip() function to each item within each list across the entire dataset. Another potential inconsistency arises from quantity information embedded within the "Product Details" column. This stage explicitly targets the pattern of "x" followed by a number, which typically indicates the quantity of a particular item purchased. This stage meticulously removes this pattern to ensure the analysis focuses solely on the type of item and not the quantity. The Python module "re" (regular expressions) comes into play here. The module allows for creating a pattern matching the format "x[number]" and its subsequent replacement with an empty string for each item within each transaction. This process removes the quantity information, leaving only the core product details for further analysis.

2.2.2 One Hot Encoding

One-hot encoding converts categorical data into numerical values [14]. For each row of data, the column representing the category to which the data belongs is filled with the value one, and the other columns are filled with the value 0.
This stage performs one-hot encoding of the 'Product Details' column based on the 'Order ID.' This process involves several steps:

a. data['Product Details'].explode(): Converts each element in the 'Product Details' column, a list of items in each transaction, into a separate row.
b. pd.get_dummies(): Returns a one-hot encoding representation of the expanded data. Each unique item in all transactions will become a new column in the result DataFrame, with a value of True indicating the presence of that item in the transaction and a value of False for items that do not exist.

2.3 Processing with FP-Growth Algorithm

Data processing is done using Google Colab with Python as the programming language. The FP-Growth Algorithm method assists in this process.
The FP-Growth (Frequent Pattern Growth) algorithm is one of the algorithms used in finding frequent itemsets (groups of items that often appear together) and association rules (relationships between those items) [15]. How FP-Growth works:

a. Prepare data: Transaction data is prepared as a list of lists, where each sub-list represents items purchased together in a single transaction.
b. Minimum support: A minimum support value is specified, indicating the minimum number of times an item appears to be considered frequent [16].

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

c. Building FP-Tree: FP-Tree data structures are built based on transaction data. FP-Tree is like a prefix tree that stores frequent itemset information efficiently.
d. Group items by frequent itemset: Using the FP-Tree structure, a recursive process is performed to find frequent itemsets based on a set minimum of support.

In FP-Growth, there is Support and Confidence. Support measures how often an item appears in all transactions in a dataset [17]. Confidence indicates how the association rules are usually correct in transactions [18]. It measures how much probability that when one set of items (antecedents) is purchased, other items (consequents) will also be purchased. Here is the equation to calculate support and confidence in Equations (1) and (2).

$$Support(X) = \frac{Number\ of\ Transactions\ Containing\ X}{Total\ Transactions} \tag{1}$$

$$Confidence(X \rightarrow Y) = \frac{Support(X \cup Y)}{Support(X)} \tag{2}$$

2.4 Data Analysis and Conclusion

This research leverages the power of market basket analysis to extract valuable insights from customer buying patterns within Muslim clothing store X. The analysis will generate frequent item sets and groups that appear together frequently in customer transactions. These frequent itemsets will then be used to identify association rules. Association rules reveal the relationships between Muslim clothing items, highlighting which items are often purchased together. The analysis will evaluate each association rule's resulting support and confidence values meticulously. Support indicates the frequency with which a particular item occurs within the dataset. At the same time, confidence reflects the likelihood of a customer purchasing a specific item, given that they have already purchased another item in the set. The study will delve deeper into customer purchase behavior by analyzing the frequent itemsets and the corresponding association rules with their support and confidence values. This comprehensive analysis will identify products with the most robust positive relationships (frequently purchased together) and those with the weakest relationships (rarely purchased together). Understanding these contrasting relationships holds significant value. Products with high positive support and confidence can inform strategic product placement decisions within the store, potentially leading to increased sales of complementary items. Conversely, identifying products with low support and confidence can highlight potential inefficiencies in product offerings, allowing store owners to optimize their inventory management strategies.

3. RESULT AND DISCUSSION

3.1 Data Preprocessing

This study employs a two-stage data preprocessing approach to ensure the data is well-suited for market basket analysis. The first stage, data cleaning, plays a vital role in guaranteeing the accuracy and integrity of the analysis. Here, meticulous efforts are undertaken to identify and rectify inconsistencies or errors within the dataset. This may involve handling missing values, removing duplicate entries, and correcting data formatting inconsistencies. Clean data forms the foundation for reliable analysis, allowing for the extraction of meaningful insights. The second stage of data preprocessing focuses on data transformation through one-hot encoding. This technique is crucial in preparing the data for the market basket analysis algorithms. One-hot encoding addresses the challenge of categorical data, where product details are represented by text labels rather than numerical values. This technique efficiently converts these categorical variables into a binary representation. Each unique product category is transformed into a separate column, and each transaction is then assigned a value of 1 for the product(s) purchased within that transaction and 0 for products not purchased. This transformation process facilitates the analysis by enabling algorithms to identify co-occurrence patterns among Muslim clothing items readily.

3.1.1 Data Cleaning

Data cleaning is identifying, deleting, or correcting inaccurate, incomplete, irrelevant, or out-of-date data from a dataset [19]. This process involves several steps, such as separating the Product Details feature into separate items, removing extra spaces or non-essential characters, and removing the "x" and the number that follows the "x" from each item in each transaction. After going through this stage, cleaner and neater data will be produced, where the Product Details feature becomes a List, making it easier to process at the One Hot Encoding stage. The results of the data-cleaning stage can be seen in Table 2.

Table 2. Clean data cleaning results

No	Order ID	Product Details
0	#19141	['Zayyan Dewasa Black L Black', 'Zayyan Dewasa Brown Olive L Brown Olive']
1	#19140	[Latif 07 Pendek Abu L Abu]
2	#19139	['Rairaka 06 Denim 10 Denim', 'Rairaka 06 Matcha 6 Matcha']
3	#19138	[Kurta Ahnaf Anak Dark Grey XXS Dark Grey]

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

4	#19137	['Kurta Ahnaf Dewasa Black S Black', 'Kurta Ahnaf Dewasa Black L Black']
...	...	...
3084	#15955	['Shaka Dewasa Black M Black', 'Shaka Dewasa Black L Black', 'Shaka …….

3.1.2 One Hot Encoding

One Hot Encoding converts each unique item in all transactions into a new field, with a value of True indicating the presence of that item in the transaction and a False value for items that do not exist [20]. One-hot encoding is a crucial data transformation technique in this study to prepare the data for the market basket analysis algorithms. This technique effectively addresses the challenge presented by categorical data, where product details are represented by text labels (e.g., "'Zayyan Dewasa Black L Black," "'Zayyan Dewasa Brown Olive L Brown Olive'") rather than numerical values. Traditional algorithms struggle to analyze such categorical data directly. One-hot encoding tackles this issue by creating a new binary representation for each unique product category within the dataset. This essentially involves transforming the original dataset with its single "Product Details" column containing text labels into a more comprehensive data frame with numerous new columns. Each new column represents a specific unique product category identified in the dataset. The values within these new columns are binary - either True or False. For a given transaction, a True value in a particular product column indicates the presence of that specific item within that transaction, while a False value signifies its absence. For instance, if the original dataset identifies 1,854 unique Muslim clothing items across all transactions, the one-hot encoding process would result in a new data frame with 1,854 additional columns (besides the original columns like Order ID), one for each unique product. In each transaction row, the corresponding product columns would be assigned a True value if that specific item was purchased and False otherwise. This transformation process effectively converts the categorical product details into a numerical representation that market basket analysis algorithms can readily work with. The algorithms can reveal insights into frequently purchased combinations of Muslim clothing items by identifying co-occurrence patterns within these binary values. The results of the One Hot Encoding process can be seen in Table 3.

Table 3. One hot encoding result data

Order ID	A04 HIJAU TUA ANAK L	A04 HIJAU TUA ANAK S	A04 MERBAT ANAK L	A04 NAVY ANAK S	...	rairaka 03 sage 2 sage	rairaka 03 sage 4 sage	rairaka 03 sage 6 sage	rairaka 03 sage 8 sage
0	#19141	False	False	False	False	False	False	False	False
1	#19140	False	False	False	False	False	False	False	False
2	#19139	False	False	False	False	False	False	False	False
3	#19138	False	False	False	False	False	False	False	False
4	#19137	False	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...	...	...
3084	#15955	False	False	False	False	False	False	False	False

3.2 Data Processing with FP-Growth Algorithm

3.2.1 Association Rules with Minimum Support

This section is the association analysis stage, which uses the FP-Growth algorithm to calculate support and confidence values. The FP-Growth algorithm generates association rules based on the frequency with which the itemset occurs [21]. The FP-Growth algorithm, a powerful technique designed for large datasets, extracts valuable insights from the preprocessed data. Association analysis, through FP-Growth, aims to identify frequently occurring combinations of Muslim clothing items within customer transactions. These combinations, referred to as frequent itemsets, represent patterns in customer buying behavior. The algorithm meticulously analyzes each transaction, identifying itemsets that appear together above a user-defined minimum support threshold. This threshold determines the prevalence level required for an itemset to be considered frequent. Selecting an appropriate minimum support value is crucial. In this study, the threshold is set at 0.009, signifying that an itemset must appear in at least 0.9% of all transactions to be considered frequent. This balances capturing commonly purchased combinations and avoiding overly specific, statistically insignificant itemsets. Once frequent itemsets are identified, FP-Growth utilizes them to generate association rules. These rules take the form "X -> Y," where X represents a subset of items within the frequent itemset, and Y represents the remaining item(s) in the same set. The association rule reveals a relationship between the items, highlighting the likelihood of purchasing Y, given that X has already been purchased. The strength of this relationship is measured by two key metrics: support and confidence. Support reflects the prevalence of a specific itemset within the dataset. The minimum support threshold of 0.009 ensures that only statistically significant itemsets are considered. A high support value for a particular item indicates a popular combination of Muslim clothing items frequently purchased by customers. Confidence delves deeper, focusing on the conditional probability of buying a specific item (Y) within an itemset given that another item (X) within the same set has already been purchased. For instance, a high confidence value associated with the rule "'Azmi Anak Army L Army -> 'Azmi Dewasa Army L Army" might reveal that customers who buy an Azmi Anak Army L Army are highly likely to also purchase a matching Azmi Dewasa

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

Army L Army within the same transaction. The study can uncover significant customer buying behavior trends by meticulously analyzing support and confidence values for various frequent itemsets and the corresponding association rules. Understanding these trends equips Muslim clothing store owners with valuable knowledge. Identifying frequently co-purchased items can inform strategic product placement decisions within the store, potentially leading to increased sales of complementary items. Additionally, insights from association analysis can be leveraged to develop targeted marketing campaigns that promote specific combinations of Muslim clothing items based on customer preferences. Table 4 is one itemset generated using the FP-Growth Algorithm.

Table 4. Dataset itemset FP-Growth

support	itemsets
0.016532	(B07 NAVY ANAK S)
0.016207	(Kurta Ahnaf Dewasa Black L Black)
0.016207	(Akmal Anak Putih Pendek L Putih)
0.015883	(Azmi Dewasa Magenta L Magenta)
0.015559	(Akmal Anak Putih Pendek XL Putih)
...	...
0.009076	(Akmal Dewasa Putih Panjang M Putih)

In Table 4, each row represents an itemset with columns indicating the items included in that itemset. The itemset with the highest support value is item B07 NAVY ANAK S with a value of 0.016532. The lowest support value is Akmal Dewasa Putih Panjang M Putih, which has a value of 0.009076. Data visualization for itemset frequency results with minimum support can be seen in Figure 2 below.

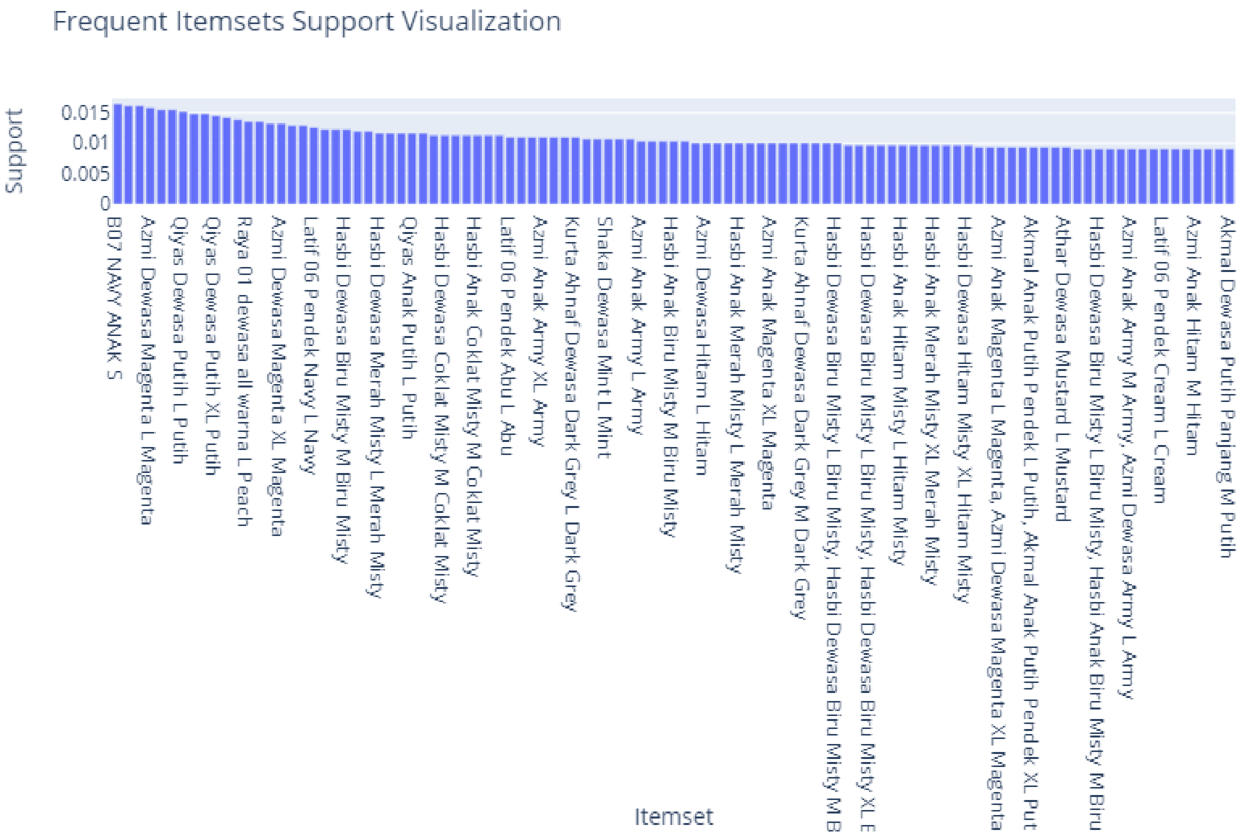

Figure 2. Itemset frequency results visualization with minimum support

In association analysis, support determines how common an itemset is in a dataset [22]. Itemsets with high support tend to be more significant in association analysis because they appear together more often with other items. If you look at Figure 2, the more to the right, the less frequently the itemset seems, and if the more to the left, the more often the itemset appears. The study delves deeper into the intricate relationships between Muslim clothing items by identifying these high-support items. It uncovers the most common pairings, trios, or even larger groupings of items that customers tend to purchase together. This knowledge empowers Muslim fashion store owners to make data-driven decisions that can significantly impact their business. For example, high-support itemsets that frequently combine can inform targeted promotions and strategic product placement within the store. By positioning these frequently co-purchased items nearby, stores can leverage the power of suggestion, influencing customer buying behavior and potentially increasing sales.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

3.2.2 Rules of Association with Minimum Confidence

This section uses the association_rules() function to find the association rules of a previously generated itemset. The metric="confidence" parameter specifies that this function uses confidence as a metric to evaluate association rules. DataFrames that contain association rules are sorted by the metric used (in this case, confidence). Each row represents a single rule of association with columns providing information about antecedents, consequents, support, confidence, lift, and several other metrics. The results of the first rule of association can be seen in Table 5.

Table 5. FP-Growth Association Rules 1

antecedents	consequents	support	confidence	lift
frozenset({'Azmi Anak Army L Army'})	frozenset({'Azmi Dewasa Army L Army'})	0,009724	0,9375	62,87364
frozenset({'Azmi Anak Magenta L Magenta'})	frozenset({'Azmi Dewasa Magenta XL Magenta'})	0,0094	0,935484	70,38946
frozenset({'Azmi Anak Army M Army'})	frozenset({'Azmi Dewasa Army L Army'})	0,009076	0,903226	60,57504
frozenset({'Hasbi Anak Biru Misty L Biru Misty'})	frozenset({'Hasbi Dewasa Biru Misty L Biru Misty'})	0,009076	0,875	56,23698
frozenset({'Hasbi Anak Biru Misty M Biru Misty'})	frozenset({'Hasbi Dewasa Biru Misty L Biru Misty'})	0,009076	0,875	56,23698
…	…	…	…	…
frozenset({'Akmal Anak Putih Pendek L Putih'})	frozenset({'Akmal Anak Putih Pendek XL Putih'})	0,0094	0,58	37,27708

In Table 5, frozenset({'Azmi Anak Army L Army'}) has a high linkage with frozenset({'Azmi Dewasa Army L Army'}); this is evidenced by a high confidence value of 0.9375, an Elevator value of 62.87364 and a support value of 0.009724., The items that have the lowest relatedness are frozenset({'Akmal Anak Putih Pendek L Putih'}) and frozenset({'Akmal Anak Putih Pendek XL Putih'}), which only has a confidence value of 0.58, an elevator value of 37.27708, and a support value of 0.0094.

Table 6. FP-Growth association rules 2

antecedents	consequents	leverage	conviction	zhangs metric
frozenset({'Azmi Anak Army L Army'})	frozenset({'Azmi Dewasa Army L Army'})	0,00957	15,76143	0,99441
frozenset({'Azmi Anak Magenta L Magenta'})	frozenset({'Azmi Dewasa Magenta XL Magenta'})	0,009267	15,294	0,9958
frozenset({'Azmi Anak Army M Army'})	frozenset({'Azmi Dewasa Army L Army'})	0,008926	10,17925	0,993475
frozenset({'Hasbi Anak Biru Misty L Biru Misty'})	frozenset({'Hasbi Dewasa Biru Misty L Biru Misty'})	0,008915	7,875527	0,992513
frozenset({'Hasbi Anak Biru Misty M Biru Misty'})	frozenset({'Hasbi Dewasa Biru Misty L Biru Misty'})	0,008915	7,875527	0,992513
…	…	…	…	…
frozenset({'Akmal Anak Putih Pendek L Putih'})	frozenset({'Akmal Anak Putih Pendek XL Putih'})	0,009148	2,343907	0,989206

In Table 6, we can see the value of leverage, conviction, and Zhang's metric. These values are only additional parameters to the previous three main parameters to strengthen the association rules analysis. Leverage measures how much the frequency of observations differs from expected if antecedents and consequents are independent [23]. In Table 6, leverage frozenset({'Azmi Anak Army L Army'}) and frozenset({'Azmi Dewasa Army L Army'}) is approximately 0.00957, indicating that the joint purchase of antecedents and consequents is only 0.00957 times more than would be expected if both were independent. Then, conviction measures how dependent consequents are on antecedents [24]. In Table 6, conviction frozenset({'Azmi Anak Army L Army'}) and frozenset({'Azmi Dewasa Army L Army'}) is 15.76143, indicating that consequents are 15.76143 times more likely to occur when antecedents are also purchased than when they are not. Zhang's metrics are an alternative measure that includes dependency and lift, focusing on the correlation between antecedents and consequents [25]. Zhang's metrics delve deeper by incorporating the concept of dependency, which goes beyond simply measuring the likelihood of one item appearing with another. It specifically considers how the presence (or absence) of the antecedent (the "if" side of the rule) influences the likelihood of finding the consequent (the "then" side). By incorporating this additional layer, Zhang's metrics provide a more comprehensive picture of the correlation between the antecedent and consequent. A positive Zhang metric value indicates a positive dependency, meaning the antecedent's presence increases the consequent's likelihood. Conversely, a negative value suggests a negative dependency, where the absence of the antecedent is associated with

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

a higher likelihood of the consequent. A value around zero implies little to no dependency between the items. Zhang's metrics offer a valuable tool for association rule mining by providing a more intricate understanding of the relationships between items in a dataset. This allows researchers and analysts to identify frequent co-occurrences and the underlying dependencies that drive those patterns. The visualization of the results of the association rule with minimum confidence can be seen in Figure 3.

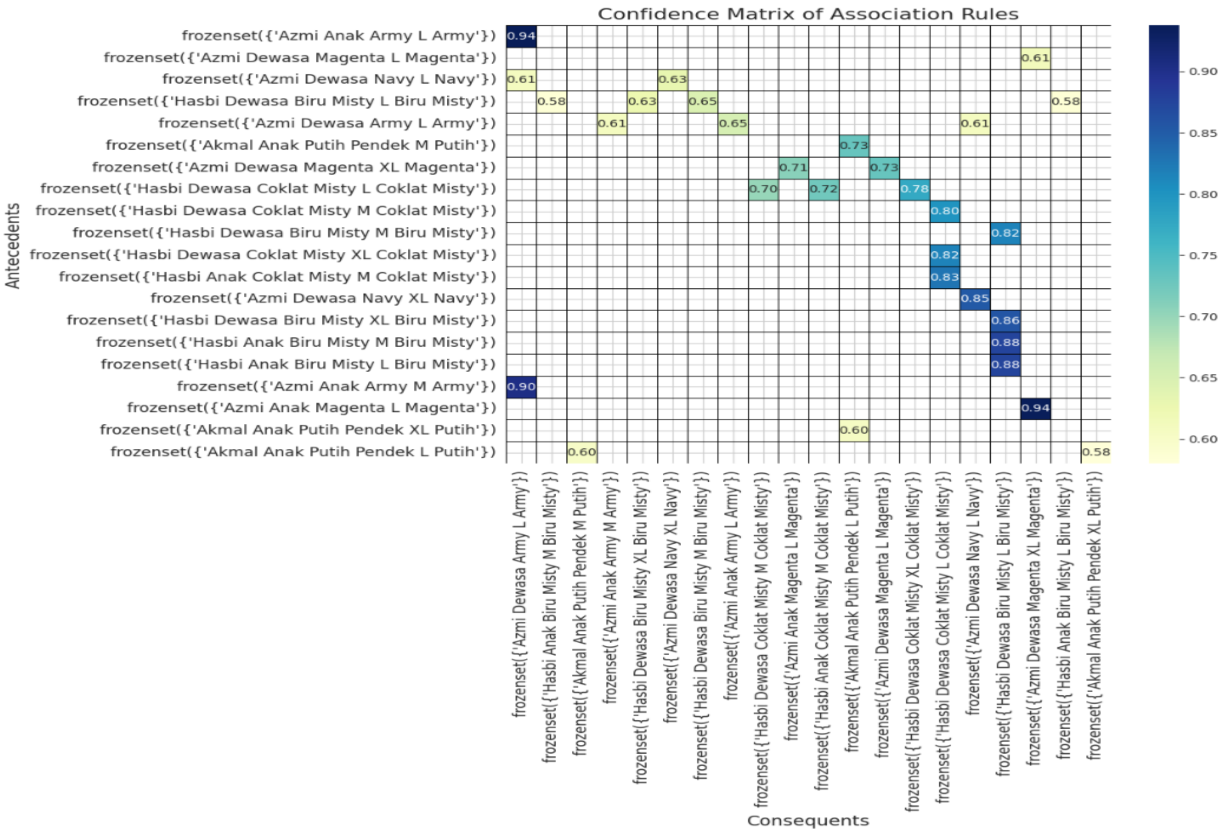

Figure 3. Visualization of results association rules with minimum confidence

4. CONCLUSION

The Muslim clothing market in Indonesia thrives on a unique blend of tradition, cultural significance, and seasonal trends. As Eid al-Fitr approaches, marking the end of Ramadan, customer buying behavior takes center stage, influencing sales strategies and inventory management for Muslim fashion businesses. This study delves into the effectiveness of market basket analysis (MBA) in understanding these customer preferences and optimizing sales during this crucial period. The study leverages the FP-Growth algorithm, a powerful technique for analyzing large datasets like the 12-month transaction record obtained from Muslim clothing store X. By applying MBA with a minimum support threshold of 0.009 and a confidence level of 0.58, the analysis aimed to uncover frequently co-purchased Muslim clothing items. This minimum support threshold ensures that only statistically significant combinations, representing at least 0.9% of transactions, are considered. The confidence level of 0.58 indicates that there's a strong likelihood (almost 60% chance) of customers purchasing a specific item (Y) within an itemset given that another item (X) from the same set has already been purchased. The analysis yielded rich insights, revealing 30 significant itemset patterns. These patterns shed light on the most compelling trend – the rising popularity of Muslim clothing designed with the concept of family and couple. This signifies a significant shift in consumer preference, with a growing demand for coordinated outfits that cater to families and couples celebrating Eid al-Fitr together. This trend presents a golden opportunity for Muslim fashion businesses to capitalize on by strategically expanding their product offerings. By integrating these data-driven insights, Muslim clothing stores can make informed decisions that maximize sales and customer satisfaction. This comprehensive approach encompasses several key areas. First, product lines can be strategically expanded to cater to the growing demand for family and couple-oriented Muslim clothing. This could involve offering matching sets for parents and children or coordinated designs for spouses. Second, inventory management can be optimized by allocating stock based on the identified frequently co-purchased items. Focusing on in-demand family and couple-oriented clothing sets ensures stores are well-prepared to meet customer needs during Eid al-Fitr. Finally, marketing and promotional campaigns can be tailored to highlight the family and couple-oriented clothing options. This could involve showcasing coordinated outfits, in-store displays, and online marketing materials or offering special promotions for family purchases. Muslim clothing stores can capitalize

Building of Informatics, Technology and Science (BITS)
Volume 6, No 1, June 2024 Page: 21−29
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i1.5162

on the flourishing family fashion trend and achieve long-term success by implementing these data-driven strategies across product selection, stock allocation, and promotions. By adopting this data-driven approach, Muslim fashion stores can capitalize on the flourishing family fashion trend during Eid al-Fitr and potentially position themselves for continued success throughout the year. As customer preferences evolve, continuous investment in customer behavior analysis using potentially even more advanced algorithms can provide deeper insights. Future research could explore these possibilities, employing these advanced algorithms for comparative analysis on even larger datasets. This ongoing exploration of customer behavior will ensure that Muslim clothing businesses remain at the forefront of this dynamic market.

REFERENCES

[1] S. Lamrhari, H. El Ghazi, M. Oubrich, and A. El Faker, “A social CRM analytic framework for improving customer retention, acquisition, and conversion,” Technol Forecast Soc Change, vol. 174, p. 121275, 2022.

[2] X. Shu and Y. Ye, “Knowledge Discovery: Methods from data mining and machine learning,” Soc Sci Res, vol. 110, p. 102817, 2023.

[3] L. Samboteng, R. Rulinawaty, M. R. Kasmad, M. Basit, and R. Rahim, “Market basket analysis of administrative patterns data of consumer purchases using data mining technology,” Journal of Applied Engineering Science, vol. 20, no. 2, pp. 339–345, 2022.

[4] A. Setiawan and R. Mulyanti, “Market Basket Analysis dengan Algoritma Apriori pada Ecommerce Toko Busana Muslim Trendy,” JUITA: Jurnal Informatika, vol. 8, no. 1, p. 11, May 2020, doi: 10.30595/juita.v8i1.4550.

[5] A. F. Lestari and M. Hafiz, “Penerapan Algoritma Apriori pada Data Penjualan Barbar Warehouse,” INOVTEK Polbeng - Seri Informatika, vol. 5, no. 1, p. 96, Jun. 2020, doi: 10.35314/isi.v5i1.1317.

[6] G. J. Pabutungan and H. D. Purnomo, “Analisa Market Basket Analysis untuk Melihat Pola Transaksi Customer Menggunakan Algoritma Apriori dan FP-Growth,” Jurnal Media Informatika Budidarma, vol. 7, no. 3, pp. 966–974, 2023.

[7] A. Wilrose, M. Afdal, S. Monalisa, and M. R. Munzir, “Penerapan Algoritma FP-Growth untuk Menentukan Strategi Promosi Berdasarkan Waktu dan Pembelian Produk,” Building of Informatics, Technology and Science (BITS), vol. 5, no. 1, Jun. 2023, doi: 10.47065/bits.v5i1.3577.

[8] R. Noviana, A. Hermawan, and D. Avianto, “Market Basket Analysis Menggunakan Algoritma Apriori dan FP Growth untuk Menentukan Pola Pembelian Konsumen,” Jurnal Media Informatika Budidarma, vol. 7, no. 3, pp. 1474–1482, Jul. 2023.

[9] A. A. Aldino, E. D. Pratiwi, Setiawansyah, S. Sintaro, and A. Dwi Putra, “Comparison Of Market Basket Analysis To Determine Consumer Purchasing Patterns Using Fp-Growth And Apriori Algorithm,” in 2021 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), IEEE, Oct. 2021, pp. 29–34. doi: 10.1109/ICOMITEE53461.2021.9650317.

[10] D. Alcan, K. Ozdemir, B. Ozkan, A. Y. Mucan, and T. Ozcan, “A Comparative Analysis of Apriori and FP-Growth Algorithms for Market Basket Analysis Using Multi-level Association Rule Mining,” in Global Joint Conference on Industrial Engineering and Its Application Areas, 2023, pp. 128–137. doi: 10.1007/978-3-031-25847-3_13.

[11] L. Shabtay, P. Fournier-Viger, R. Yaari, and I. Dattner, “A guided FP-Growth algorithm for mining multitude-targeted item-sets and class association rules in imbalanced data,” Inf Sci (N Y), vol. 553, pp. 353–375, 2021.

[12] M. Shawkat, M. Badawi, S. El-ghamrawy, R. Arnous, and A. El-desoky, “An optimized FP-growth algorithm for discovery of association rules,” J Supercomput, vol. 78, no. 4, pp. 5479–5506, 2022.

[13] J. P. B. Saputra, S. A. Rahayu, and T. Hariguna, “Market Basket Analysis Using FP-Growth Algorithm to Design Marketing Strategy by Determining Consumer Purchasing Patterns,” Journal of Applied Data Sciences, vol. 4, no. 1, pp. 38–49, Jan. 2023, doi: 10.47738/jads.v4i1.83.

[14] M. K. Dahouda and I. Joe, “A deep-learned embedding technique for categorical features encoding,” IEEE Access, vol. 9, pp. 114381–114391, 2021.

[15] I. Riadi, H. Herman, F. Fitriah, S. Suprihatin, A. Muis, and M. Yunus, “Implementation of association rule using apriori algorithm and frequent pattern growth for inventory control,” Jurnal Infotel, vol. 15, no. 4, pp. 369–378, 2023.

[16] M. Ghosh, A. Roy, P. Sil, and K. C. Mondal, “Frequent itemset mining using FP-tree: a CLA-based approach and its extended application in biodiversity data,” Innov Syst Softw Eng, vol. 19, no. 3, pp. 283–301, 2023.

[17] M. Sadeequllah, A. Rauf, and N. Alnazzawi, “Probabilistic Support Prediction: Fast frequent itemset mining in dense data,” IEEE Access, 2024.

[18] T. Osadchiy, I. Poliakov, P. Olivier, M. Rowland, and E. Foster, “Recommender system based on pairwise association rules,” Expert Syst Appl, vol. 115, pp. 535–542, 2019.

[19] F. Bukhatwa, A. Laarfi, and I. Salem, “Dirty Data between Errors and Their Handling—A Firsthand Experience in Solving Dirty Data from Within,” Int J Intell Sci, vol. 13, no. 2, pp. 48–62, 2023.

[20] A. Alazeb, “Utilizing Machine Learning with Unique Pentaplet Data Structure to Enhance Data Integrity.,” Computers, Materials & Continua, vol. 77, no. 3, 2023.

[21] H.-J. Jang, Y. Yang, J. S. Park, and B. Kim, “FP-growth algorithm for discovering region-based association rule in the IoT environment,” Electronics (Basel), vol. 10, no. 24, p. 3091, 2021.

[22] E. Hikmawati, N. U. Maulidevi, and K. Surendro, “Minimum threshold determination method based on dataset characteristics in association rule mining,” J Big Data, vol. 8, pp. 1–17, 2021.

[23] A. Upadhyaya, M. K. Mishra, and A. Saxena, “User Preferences for AI-based Healthcare Apps: an Association Mining Analysis,” SN Comput Sci, vol. 5, no. 5, p. 464, 2024.

[24] A. Z. Philipp-Muller, L. E. Wallace, and D. T. Wegener, “Where does moral conviction fit?: A factor analytic approach examining antecedents to attitude strength,” J Exp Soc Psychol, vol. 86, p. 103900, 2020.

[25] C. C. J. Hryhoruk, C. K. Leung, J. Li, B. A. Narine, and F. Wedel, “Multi-level Frequent Pattern Mining on Pipeline Incident Data,” in International Conference on Advanced Information Networking and Applications, Springer, 2024, pp. 380–392.