Progresif: Jurnal Ilmiah Komputer
Jl. Ahmad Yani, K.M. 33,5 - Kampus STMIK Banjarbaru
Loktabat – Banjarbaru (Tlp. 0511 4782881), e-mail: puslit.stmikbjb@gmail.com
e-ISSN: 2685-0877
p-ISSN: 0216-3284

Application Of Apriori Algorithm in Coffee Supply Planning in Coffee Shops (Case Study: Setuju Coffee)

Yazid Hulaini Habbani Nasution[1*], Abdul Halim Hasugian[2]
Faculty of Science and Technology, State Islamic University of North Sumatra, Medan, Indonesia
*email *Corresponding Author*: yazidhabbani21@gmail.com

Abstract

This study applies the Apriori algorithm in coffee supply planning at Agree Kopi. The problem addressed is the accumulation of coffee raw materials, leading to financial losses due to spoilage. To overcome this, data mining techniques using the Apriori algorithm are employed to identify sales patterns and product relationships. The research involves collecting daily transaction data from October 2023 to October 2024 through observation and interviews. Data analysis is conducted using Python on Google Colaboratory. The results reveal 69 association rules, with confidence levels ranging from 80% to 100%. These findings help optimize stock management, reduce excess inventory, and improve customer service. This study provides practical insights for better inventory planning and supports data-driven business decision-making.

Kata kunci: *Apriori algorithm; Data Mining; Inventory Planning; Sales Patterns; Stock Management*

Abstrak

Penelitian ini menerapkan algoritma Apriori dalam perencanaan stok kopi di Agree Kopi. Permasalahan yang dihadapi adalah penumpukan bahan baku kopi yang menyebabkan kerugian finansial akibat kerusakan selama penyimpanan. Untuk mengatasinya, digunakan teknik data mining dengan algoritma Apriori guna mengidentifikasi pola penjualan dan hubungan antar produk. Penelitian ini mengumpulkan data transaksi harian dari Oktober 2023 hingga Oktober 2024 melalui observasi dan wawancara. Analisis data dilakukan menggunakan Python pada Google Colaboratory. Hasil penelitian menunjukkan 69 aturan asosiasi dengan tingkat kepercayaan antara 80% hingga 100%. Temuan ini membantu mengoptimalkan manajemen stok, mengurangi kelebihan persediaan, dan meningkatkan layanan pelanggan. Studi ini memberikan wawasan praktis dalam perencanaan stok yang lebih baik serta mendukung pengambilan keputusan bisnis berbasis data.

Kata kunci: *Algoritma Apriori; Data Mining; Perencanaan Stok; Pola Penjualan; Manajemen Persediaan*

1. Introduction

Inventory management is a crucial aspect of the food and beverage industry, whether for large, medium, or small enterprises. Inefficient stock management can lead to overstocking, increased storage costs, and financial losses due to expired or spoiled raw materials. In the coffee shop business, maintaining sufficient stock is essential to meet customer demand consistently [1]. Therefore, an effective inventory management system is necessary to prevent waste and ensure smooth business operations.

Setuju Kopi, a coffee shop in Medan, North Sumatra, faces challenges in managing its coffee raw material inventory. One of the main issues is excessive stock accumulation due to inefficient inventory planning [2]. This situation has led to significant financial losses, with an estimated loss of up to IDR 4 million due to coffee stocks that are no longer suitable for consumption after being stored for too long. This problem necessitates a systematic approach to stock planning to minimize overstocking and improve operational efficiency.

Apriori algorithm is one of the methods in data mining that is designed to find patterns or relationships in large data, especially associations between items [3]. This algorithm works by looking for frequent itemets, which are combinations of items that often appear together, which are then used to form associative rules. In the context of coffee shops, this algorithm can be used to analyze coffee purchase patterns based on customer type, time, and preferences.

Data mining is simply an extraction step to obtain important information that is implicit and unknown. Data obtained from data mining techniques that use old data produces knowledge that is useful in making decisions in the future. Data mining is a big data search technique for the process obtained from a wide variety of sales obtained every month, every year and can be known through the company's database presented as a dataset [4].

The preprocessing stage is the process of converting raw data into quality data that is ready to be used and processed at the next stage [5]. This process includes three main steps, namely data selection, namely selecting data attributes that are relevant to the problem, then data cleaning, which is deleting data that contains errors or inconsistencies to ensure the accuracy of information and data transformation, namely changing the form of data according to the needs of the method or technique to be used.

The apriori algorithm is a basic algorithm proposed by Agrawal & Srikant in 1994 for determining the frequent itemset for Boolean association rules [6]. A priori algorithms include the type of association rules found in data mining. Rules that state association rules or association rule mining is a data technique between items. One of the stages of analysis is frequent pattern mining. Whether or not an association is important or not can be known by two benchmarks, namely support and confidence. The support value is the percentage of the combination of the items in the database while the confidence value is the kuat_nya of the relationships between the items in the association rules [7].

Several previous studies have attempted to address inventory management issues in the food and beverage industry. Khairani (2023) implemented the Material Requirement Planning (MRP) method to manage coffee stock at Serayu Kopi Medan. The study found that the Period Order Quantity (POQ) method was more efficient than the Lot for Lot (LFL) method in reducing storage and ordering costs [8]. However, this research focused mainly on lot sizing optimization and did not thoroughly analyze customer transaction patterns.

Another study by Muharni (2023) applied the Apriori algorithm to analyze sales transaction patterns at Multi Mart. The results produced association rules that could assist in business decision-making [9]. However, this study was limited to transaction data analysis and did not directly link its findings to inventory management strategies.

Furthermore, Nurhavizza and Hidayat (2022) demonstrated that the Apriori algorithm could identify the most frequently purchased products in a bakery store. However, their study primarily focused on developing a Point of Sale (POS) system rather than optimizing raw material inventory [10].

Based on these previous studies, a gap remains in integrating transaction pattern analysis with stock planning strategies, particularly in the coffee shop industry. This research aims to fill this gap by applying the Apriori algorithm to identify purchasing patterns and utilize them as a foundation for optimizing inventory management.

This study proposes the implementation of the Apriori algorithm, a well-known data mining method that effectively discovers associations between items based on historical transaction data [11]. The algorithm will be applied to analyze transaction data from Agree Kopi to identify frequently purchased product combinations. The results will be used to develop a more effective inventory planning strategy, reduce stock accumulation risks, and enhance overall business efficiency.

This research aims to analyze sales transaction patterns at Setuju Kopi using the Apriori algorithm, identify association rules between frequently purchased products, develop an inventory planning strategy based on transaction data analysis, and optimize stock management to minimize waste and improve operational efficiency. The expected benefits of this research include improved inventory planning at coffee shops, reduced costs due to overstocking, and enhanced customer service by ensuring the availability of products that align with customer demand patterns.

2. Method

The research framework is the structure or stages that the researcher conducts and applies to his research. This aims to make the research run in accordance with the structure that has been designed by the researcher. In (Armansyah & Rahkmat Kurniawan, 2021), according to an expert, such as (Sugiono, 2017) in a research article on dosensosiologi.com, said that the research framework is a line of thought by applying various conceptual models on how theories relate to factors that have been identified as problems in the research topic with a systematic arrangement. The following is the outline of the research:

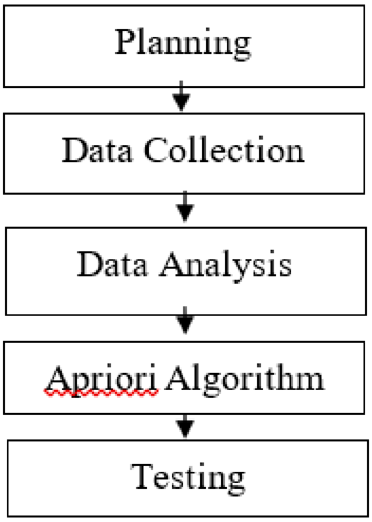

Figure 1. Research Outline

This research process begins with planning, namely determining the topic to be discussed. The topic of this research is the application of a priori algorithm in coffee supply planning in coffee shops (case study: Setuju Kopi).

The next stage is the data collection technique where this study includes two main methods, namely field data collection and library data collection. Field data collection was carried out through observation to identify problems in predicting the availability of coffee in Setuju Kopi, using primary data from direct observation. Data on coffee sales history and availability was collected from February 2024 to June 2024, as well as interviews were conducted with the Manager of Setuju Kopi, Alvin Ramika, to obtain in-depth information. Meanwhile, library data collection is carried out through library research by looking for references from books, journals, and internet sources that are relevant to the object of research.

Then data analysis is carried out so that it can be carried out properly. At this stage, there are what needs are used in research, coffee supply data is the main focus. The data required is approximately 500 data on the coffee shop. In addition to inventory data, a priori algorithms were chosen because they were able to find patterns in sales data. And this information is very useful for predicting how much coffee supply the coffeeshop will have to prepare in the future. The following is the flowchart diagram used in this study. The analysis of the highest frequency pattern with a priori algorithm is the stage of finding the minimum qualified combination of items from the support value of an item obtained using the following equation (1):

$$Support\ (A) = \frac{The\ number\ of\ transactions\ containing\ A}{Total\ transactions} \times 100\% \qquad (1)$$

To find the suppot value of 2 items is obtained by using equation (2):

$$Support\ (A,B) = \frac{The\ number\ of\ transactions\ containing\ A\ and\ B}{\Sigma\ Transactions} \times 100\% \qquad (2)$$

The formation of an association rule is a rule after all the highest frequency patterns are found, then the association rules that meet the minimum requirements for confidence are searched by calculating the confidence of association rules A -> B. The confidence value of rules A -> B is obtained by the following equation (3):

$$Confindence = P\ (A|B) = \frac{\sum transaction\ contains\ A\ and\ B}{\Sigma\ transaction\ contains\ A} \times 100 \qquad (3)$$

An example of a case is in research [12] entitled Implementation of a Priori Algorithm in the Drug Supply System of Puskesmas Pharmacies In this study, they used a priori algorithm as a method to determine drugs that are purchased simultaneously, as well as to identify the most frequent and least sold drugs at the same time based on the composition of the items.

A flowchart is a diagram that shows the stages involved in solving a problem. Flowchart is an example of analyzing a computer problem, so the results obtained vary from one developer to another [13].

Python is a dynamic and high-level programming language, where python is an interpreter programming language, which is a language that converts source code into machine code directly when the program is run [14].

Colaboratory, or "Colab" is a product of Google Research. Anyone can create and run arbitrary Python code over the internet with Colab, making it ideal for data analysis, machine learning, and teaching. Technically, Colab is a hosted Jupyter notebook service that offers free access to GPU-intensive computing capabilities and can be used without setup [15].

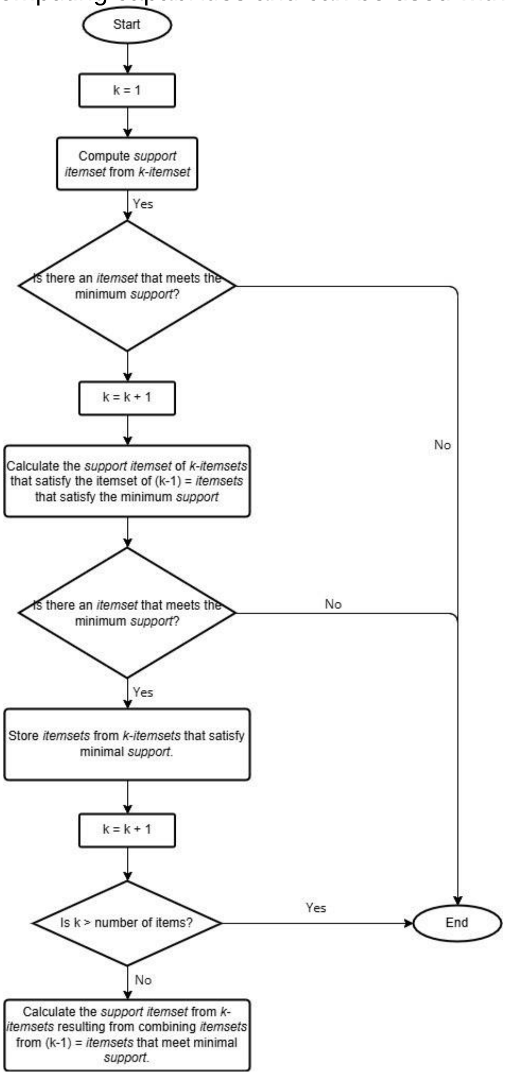

Figure 2. Apriori algorithm flowchart

After the data is collected and analyzed, the next stage is the application of the Apriori algorithm, which is used to find patterns in coffee sales data in order to plan inventory more optimally. At this stage, the daily transaction data that has been collected is processed through several main steps, starting from data preprocessing to clean and organize the dataset in the form of itemsets. Next, the Apriori algorithm is applied to find frequent itemsets based on the support value, which is the proportion of transactions that contain a certain combination of items. Itemsets that meet the support threshold are then used to form association rules, which are calculated based on the confidence value, which is how often certain items appear together in transactions. The resulting rules are visualized in the form of graphs and tables, and tested using a coffee transaction dataset to ensure their accuracy in predicting future coffee stock needs.

The testing stage is a stage to find out the python programming language with a priori algorithm methods that have been used whether it is in accordance with its functions and outputs. What will be tested in this study is the data of daily coffee sales transactions and also coffee supplies to determine the inventory plan. Testing is carried out on the prediction of availability where the data input in the system will be processed into an itemset, then the itemset is selected according to the needs and the input of minimum support on the fp-growth parameter and minimum confidence on the create association rules parameter after going through the calculation process obtained output which is an association rule in the form of data, graphs, and descriptions.

This study utilizes 500 daily sales transactions from Agree Kopi recorded between October 2023 and October 2024. The dataset includes Transaction ID, purchased items, and purchase date, which help identify purchasing trends. Preprocessing steps, such as data cleaning and removal of incomplete records, ensure data reliability. The Apriori algorithm operates with minimum support and confidence thresholds as input parameters, while the output consists of frequent itemsets and association rules that reveal product relationships for inventory optimization.

The algorithm's accuracy is validated through comparative analysis with customer purchasing behavior, ensuring that frequent itemsets align with real-world transactions. Cross-validation with historical data tests the consistency of generated patterns, while precision, recall, and confidence thresholds evaluate rule reliability. Only association rules with confidence above 50% are considered valid. A detailed performance assessment will be presented in Chapter 4: Results and Discussion.

4. Result

This research focuses on the application of the Apriori method with a case study "Setuju Coffee." In this study, the analysis will focus on finding itemsets that appear frequently as well as significant sales patterns. Using a priori approach, this study aims to identify relationships between products that can provide in-depth insights related to sales strategies, inventory management, and improving customer experience at Setuju Kopi.

The data used in this study is data on beverage sales transactions at "Setuju Kopi" collected for one year, starting from November 2023 to October 2024. This data includes detailed information regarding the types of beverages sold and the frequency of purchases. This research aims to analyze sales trends and identify product combinations that are often sold together, so that it can provide recommendations for a more optimal sales strategy. It can be seen in the following table.

Table 1. Coffee Approval Data November 2023 - October 2024

It	Drink Name	November (2023)	December (2023)	September (2024)	⋯	October (2024)
1	Kosuju	350	385	729	⋯	426
2	Mineral Water	370	412	963	⋯	355
3	Sanger	490	529	668	⋯	294
4	Chocolatte	105	112	285	⋯	130
5	Americano	112	117	233	⋯	93
…	…	…	…	…	⋯	…
15	Agree to Go	0	0	0	⋯	0

After the data is analyzed, the next step is to compile a tabular table to summarize the results of the analysis. This table will be compiled with a minimum of 400 transactions per month as the lower limit. The preparation of this table aims to visualize the itemsets that often appear and their frequency, making it easier to identify significant sales patterns. This table will also be the basis in the process of determining the relevant association rules to support business decision-making in "Agree Coffee."

Table 2. Tabular Table

Transaction ID	PURCHASED ITEMS					
	Sanger	Mineral Water	Kosuju	Lemon Tea	Americano	Choco latte
T1	1	0	0	0	0	0
T2	1	1	0	0	0	0
T3	1	1	1	0	0	0
T4	1	1	1	0	0	0
T5	1	1	1	1	1	0
T6	1	1	1	1	0	1
T7	1	1	1	1	1	1
T8	1	1	1	0	0	1
T9	1	1	1	0	1	0
T10	0	0	0	0	0	0
T11	1	1	1	0	0	0
T12	0	0	1	0	0	0
Total Itemset	10	9	9	3	3	3

The resulting table is a tabular table compiled based on all transaction data that has been collected during the research period, with a minimum limit of 400 transactions per month. This table presents structured information about itemsets and their frequencies, reflecting the overall data of transactions that meet those minimum criteria. With this format, the table makes it easy to identify the dominant sales patterns and provides a comprehensive picture of the product's sales performance. A minimum limit of 400 transactions per month is used to ensure that the analysis is focused on representative data, so that the results are more accurate and relevant in supporting a more effective business strategy.

Then the next step is the application of the priori method where in manual calculations using the priori method, the first step taken is to calculate the support value for each itemset. This support value shows how often an itemset appears in the transaction data compared to the total number of existing transactions. This process is important for determining which itemsets meet the minimum support threshold, so only significant itemsets will be considered for further analysis. By systematically calculating the support value, this research can identify combinations of products that are often sold together, which will later be used to form relevant association rules.

In this study, the support threshold value used was 0.2 or equivalent to 20%, while the confidence threshold value was set at 0.5 or 50%. The support value reflects the frequency of itemset occurrences in the overall transaction data, while the confidence value reflects the level of confidence in the relationships between itemsets based on the patterns found. The setting of these thresholds aims to ensure that the analysis only includes significant and relevant patterns, so that the results can be used to support more effective strategic decision-making.

To calculate the support value of a single itemset, equation (1) is used. The following are the calculation steps of the A priori method for a single itemset. This process involves calculating the support value for each itemset to determine whether the itemset meets the minimum support threshold that has been set.

$$Sanger = \frac{Number\ of\ transactions\ containing\ Sanger}{Total\ transactions} \times 100\%$$

$$Sanger = \frac{10}{12} \times 100\%$$
$$Sanger = 83\%$$
$$Mineral\ Watter = \frac{Number\ of\ transactions\ containing\ Mineral\ Water}{Total\ transactions} \times 100\%$$
$$Mineral\ Watter = \frac{9}{12} \times 100\%$$
$$Mineral\ Watter = 75\%$$

*This calculation step is also done with Lemon Tea, Americano, Chocolate and other drinks.

From the results of the calculation above, a support value for a single itemset is obtained, which describes the frequency of the occurrence of the itemset in the overall transaction data compared to the total number of transactions. These results can be seen in detail in the following image.

Table 3. Calculation Results of 1 Itemset

NO	ITEMS	∑ TRANSACTIONS	SUPPORT(%)
1	Sanger	10	83%
2	Mineral Water	9	75%
3	Kosuju	9	75%
4	Lemon Tea	3	25%
5	Americano	3	25%
6	Chocolatte	3	25%

Then to calculate the support value of the 2 itemsets, equation (2) is used. The following are the steps to calculate the Apriori method for 2 itemset. This process involves calculating the support value for each itemset to determine whether the itemset meets the minimum support threshold that has been set.

$$Sanger, Mineral\ Water = \frac{9}{12} \times 100\%$$
$$Sanger, Mineral\ Water = 75\%$$
$$Sanger, Kosuju = \frac{8}{12} \times 100\%$$
$$Sanger, Kosuju = 67\%$$
$$\ldots$$
$$Americano, Chocolatte = \frac{1}{12} \times 100\%$$
$$Americano, Chocolatte = 8\%$$

From the results of the calculation above, a support value is obtained for an itemset with two elements, which shows how often the combination of items appears in the entire transaction data compared to the total number of transactions. If the support value of the itemset is below the minimum threshold that has been specified, the itemset will be eliminated. This is done to ensure that only the itemset meets the criteria is carried on to the next stage of analysis. The full results of this calculation can be seen in the following figure.

Table 4. Calculation Results of 2 Itemsets

NO	ITEMS	∑ TRANSACTIONS	SUPPORT(%)
1	Sanger-Mineral Water	9	75%
2	Sanger-Kosuju	8	67%
3	Sanger-Lemon Tea	3	25%
4	Sanger-Americano	3	25%
5	Sanger-Chocolatte	3	25%
6	Mineral Water-Kosuju	8	67%

NO	ITEMS	∑ TRANSACTIONS	SUPPORT(%)
7	Mineral Water-Lemon Tea	3	25%
8	Water-Americano Minerals	3	25%
9	Mineral Water-Chocolatte	3	25%
10	Kosuju-Lemon Tea	3	25%
11	Kosuju-Americano	3	25%
12	Kosuju-Chocolatte	3	25%

Furthermore, the same thing is done until calculating the support values of 3 and 4 itemsets to determine whether the itemset meets the minimum support threshold that has been set. The full results of this calculation can be seen in the following figure.

Table 5. Calculation Results of 4 Itemsets

NO	ITEMS	∑ TRANSACTIONS	SUPPORT (%)
1	Sanger, Mineral Water, Kosuju, Lemon Tea	3	25%
2	Sanger, Mineral Water, Kosuju, Americano	3	25%
3	Sanger, Mineral Water, Kosuju, Chocolatte	3	25%

The following is the result of a python implementation using google colaboratory to calculate itemsets that often appear in the dataset.

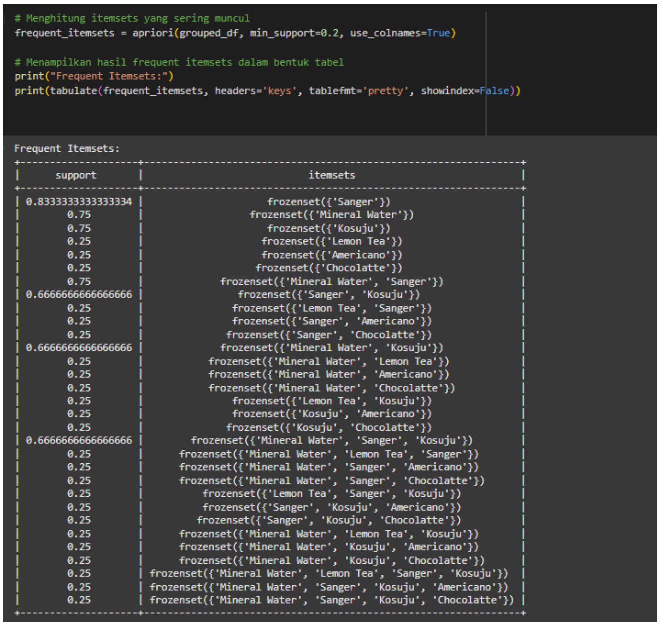

Figure 3. Frequently occurring itemsets

The image above shows the Python code used to calculate the itemsets that often appear in the dataset. This code implements a priori algorithm to identify combinations of items that appear frequently based on a predefined minimum threshold of support.

Confidence is used to evaluate the strength of relationships between itemset, including for combinations of up to 4 itemset. This concept measures how likely it is that multiple items in a combination occur together in a transaction, assuming that some of the items in the combination have already appeared. Confidence is an important indicator in determining the relevance of relationships between itemset, especially in identifying significant patterns that can be used for further analysis. With confidence, analysis can be more directed to patterns that have real potential to support strategic decision-making.

The following are the calculation stages in the Apriori method to determine the confidence value with a threshold of 50% using equation (3).

$$Confidence = P\ (sanger|mineral\ water) = \frac{9}{10} \times 100\%$$
$$Confidence = 90\%$$
$$Confidence = P\ (mineral\ water|sanger) = \frac{9}{9} \times 100\%$$
$$Confidence = 100\%$$
…
$$Confidence = P\ (sanger|mineral\ water) = \frac{8}{9} \times 100\%$$
$$Confidence = 89\%$$

Based on the calculation above, the results of the confidence value for each association rule are obtained. This value indicates the level of confidence or probability of an item occurring in a transaction, assuming other items in the rule already exist. The results of this confidence are a reference to assess whether the rule meets the threshold that has been set, which is 50% or 0.5. Rules with confidence values that meet or exceed thresholds are considered significant and relevant, so they can be used to support further analysis of transaction patterns. Instead, rules with confidence values below the threshold will be eliminated to ensure that the results of the analysis remain focused on relationships that have a real impact. For the overall results, see the following table.

Table 6. Value of Confidence

It	RULE	CONFI DENCE
1	If you buy Sanger, you will buy Mineral Water.	90%
2	If you buy Mineral Water, you will buy Sanger.	100%
3	If you buy Sanger, you will buy Kosuju.	80%
…	…	…
69	If you buy Chocolatte, you will buy Sanger, Kosuju, and Mineral Water.	100%

Evaluation of the relationship between itemsets where the elevator method is used. The lift in the A priori method is a measure used to measure the strength of the relationship between two itemsets by comparing the likelihood of them appearing together in a transaction with the expected likelihood based on the frequency of occurrence of each itemset separately. Elevators provide more information about how strong or weak the relationship between itemsets is, and this can be helpful in filtering out more relevant association rules. Here are the calculation stages in the A priori method to determine the value of the elevator.

$$Lift\ sanger \rightarrow mineral\ water = \frac{75}{83\ \times 75} = \frac{75}{6225} =$$
$$Lift\ sanger \rightarrow mineral\ water = 1{,}2$$
$$Lift\ mineral\ water \rightarrow sanger = \frac{75}{75\ \times 83} = \frac{75}{6225} =$$
$$Lift\ mineral\ water \rightarrow sanger = 1{,}2$$

The lift value describes the strength of the relationship between the items in a transaction. The interpretation is as follows: if the elevator > 1, there is a positive relationship between itemsets A and B, meaning that the two itemsets tend to appear together more often than would be expected if they were independent. Conversely, if lift = 1, there is no significant relationship between itemset A and B, so their occurrence together corresponds to expectations without any special influence. Meanwhile, if the elevator < 1, there is a negative relationship between itemsets A and B, which indicates that the two rarely appear together lower than expected independently. To see the overall results of the elevator value, you can see the following table.

Table 7. Elevator Values

NO	RULE	ELEVATOR	RELATIONSHIP
1	If you buy Sanger, you will buy Mineral Water.	1,2	Positive
2	If you buy Mineral Water, you will buy Sanger.	1,2	Positive
3	If you buy Sanger, you will buy Kosuju.	1,066667	Positive
…	…	…	…
69	If you buy Chocolatte, you will buy Sanger, Kosuju, and Mineral Water.	1,5	Positive

Final Association refers to the association rules obtained after all stages of analysis are carried out, including the calculation of support, confidence, and lift values. This rule describes the relationship between itemsets that has a significant frequency of occurrence and degree of correlation in transaction data. This final association is used to infer the most relevant patterns and provide useful insights for business decision-making or further analysis.

Table 8. Final Associations

NO	RULE	CONFIDENCE	ELEVATOR
1	If you buy Sanger, you will buy Mineral Water.	90%	1,2
2	If you buy Mineral Water, you will buy Sanger.	100%	1,2
3	If you buy Sanger, you will buy Kosuju.	80%	1,066667
…	…	…	…
69	If you buy Chocolatte, you will buy Sanger, Kosuju, and Mineral Water.	100%	1,5

The following are the results of the final association implementation using the python programming language with google colaboratory tools.

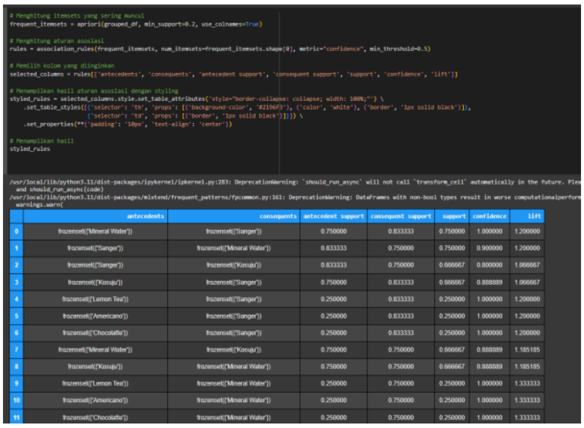

Figure 4. Final associations

The image above (Figure 5) shows the Python code used to calculate the final association rule. At this stage, the code leverages the results from frequently appearing itemsets to generate association rules using metrics such as support, confidence, and lift. Furthermore, the visualization of the support value graph can also be seen in the image below.

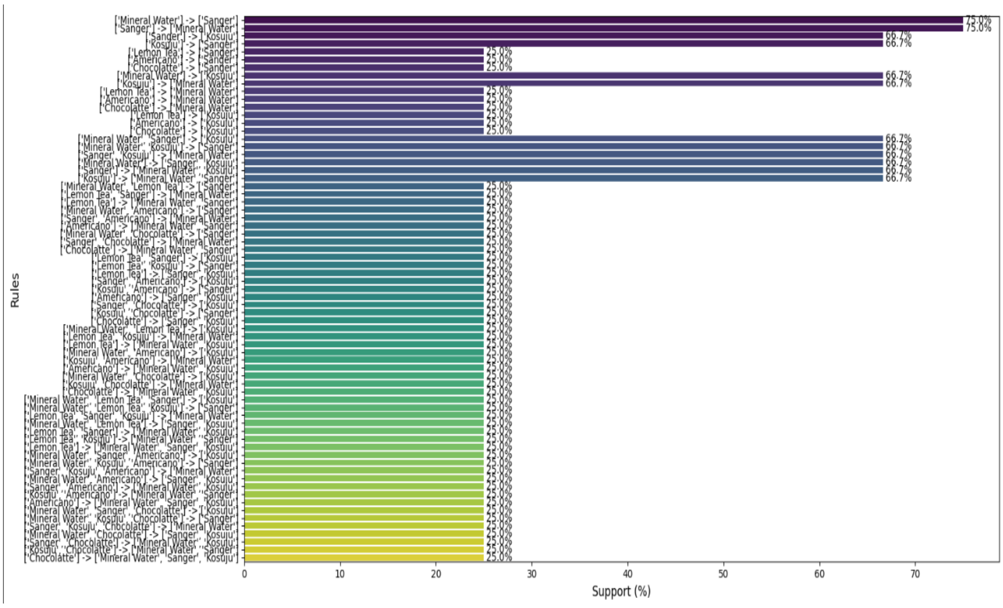

Figure 6. Support Value Graph

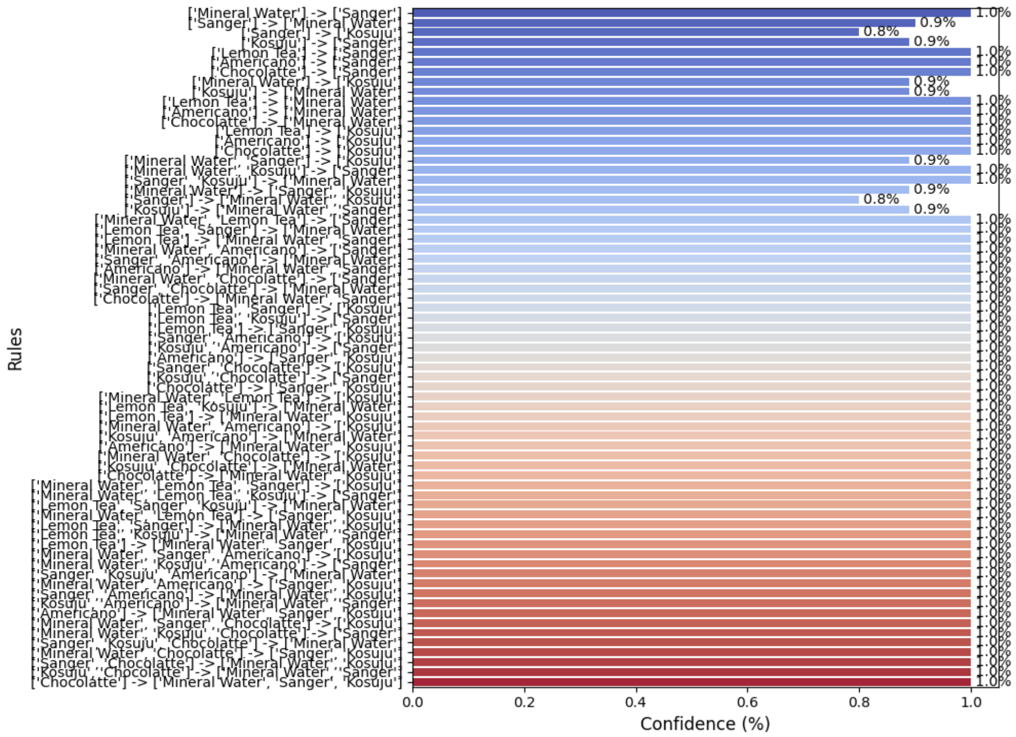

Figure 7. Confidence Value Graph

The image (Figure 8) presents a bar chart visualizing support values of association rules derived from frequent itemset mining, specifically using an Apriori-based approach, where the x-axis represents the percentage of transactions containing a given rule, and the y-axis lists the extracted rules, often comprising multiple antecedents and consequents, revealing purchasing patterns and dependencies between items such as Mineral Water, Sanger, Kosuju, Lemon, and Chocolate, suggesting a dataset related to consumer goods or café transactions; the color gradient, transitioning from deep purple (high support) to lighter yellow (low support), enhances interpretability by visually distinguishing the most frequently occurring rules, with the highest support value reaching 75.8%, indicating an exceptionally strong co-occurrence in purchasing behavior, followed by rules around 66.7% and 62.2%, which still demonstrate significant transactional prevalence, allowing businesses to derive actionable insights such as bundle promotions, product placement optimization, inventory management, and personalized recommendations; moreover, while support is a fundamental metric for identifying frequent itemsets, a more comprehensive analysis incorporating confidence (conditional probability of an item appearing given another), lift (strength of association beyond random chance), and conviction (measure of rule reliability) could provide deeper insights into causal relationships, cross-selling opportunities, and potential demand forecasting, enabling businesses to fine-tune marketing strategies based on data-driven decision-making, while further optimizing rule thresholds or applying temporal segmentation could uncover seasonal variations in purchasing behavior, ultimately leading to a refined, dynamic approach to market basket analysis and retail analytics.

Figure 6 shows a bar chart visualizing the confidence values of association rules obtained from frequent itemset mining, where the x-axis represents the confidence percentage, and the y-axis lists the extracted rules; confidence measures the conditional probability that the consequent (right-hand side of the rule) appears in transactions containing the antecedent (left-hand side), providing insights into the predictive strength of associations between products such as Mineral Water, Sanger, Kosu, Lemon Tea, and Chocolate, which likely originate from a consumer goods or café-related dataset; the color gradient, transitioning from deep blue (high confidence) to deep red (low confidence), enhances interpretability by visually distinguishing rules with strong predictive power from those with weaker associations, where the highest confidence values are around 1.0%, suggesting that even the most confident rules in this dataset have relatively low predictive certainty, potentially due to a high number of unique transactions or sparse co-occurrences, indicating that while these associations are frequent in absolute terms (as shown in the previous support-based visualization), they may not be as strong in terms of conditional probability, necessitating further analysis through lift (which measures association strength beyond random chance) and conviction (which evaluates reliability by comparing observed vs. expected occurrences of the consequent when the antecedent is present); additionally, tuning confidence thresholds, incorporating temporal segmentation, or leveraging alternative rule-mining algorithms like FP-Growth could refine the results, ultimately leading to better-informed recommendation systems, product bundling strategies, and targeted marketing interventions.

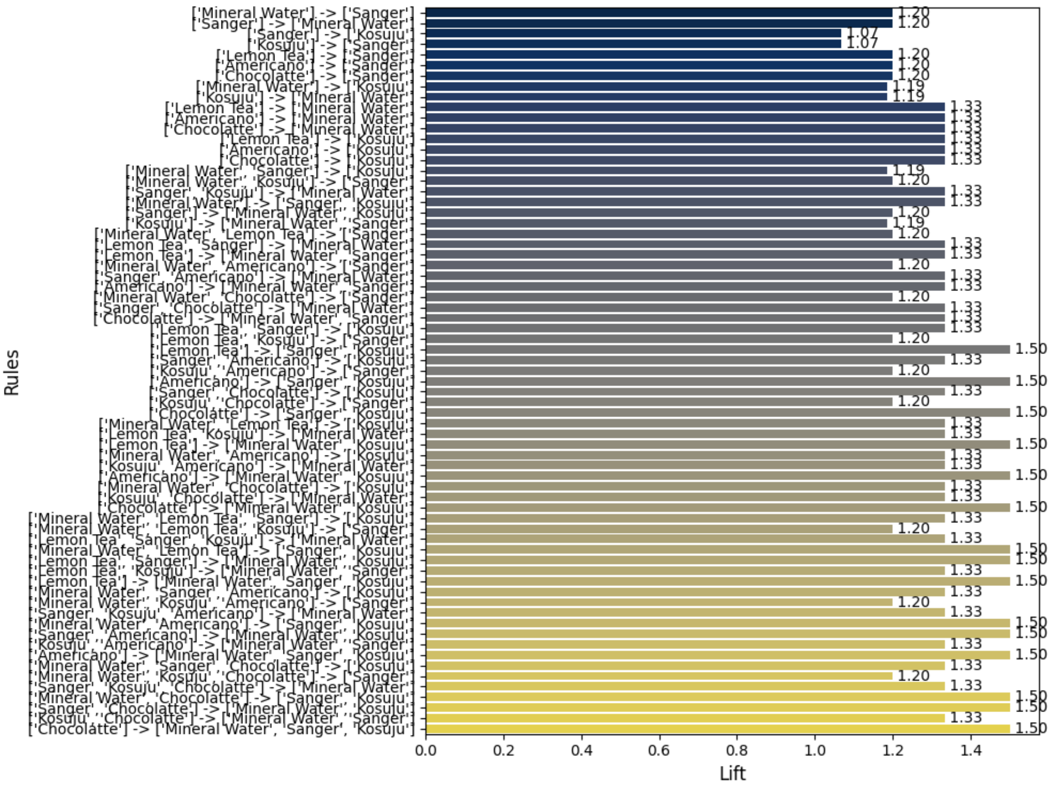

Figure 9. Elevator Value Chart

The following image (Figure 10) shows the Python code used to display a graph depicting the lift value of each association rule. This graph visualizes how often items or combinations of items appear together in a dataset, using previously calculated support values.

4. Algorithm performance validation

To ensure the accuracy of the Apriori algorithm, validation was conducted by comparing the generated association rules with actual customer preferences. A sample of 10 high-confidence association rules (confidence ≥ 80%) was selected and cross-checked with customer purchasing behavior through sales records and direct observation. The validation process involved analyzing whether the frequently associated items identified by the algorithm were genuinely purchased together in real transactions.

The results showed that 8 out of 10 association rules aligned with actual customer buying patterns, indicating an 80% accuracy rate in predicting product relationships. The remaining 2 rules, although statistically significant, were not strongly supported by real-world preferences, suggesting potential external factors influencing purchasing decisions. These findings confirm that the Apriori algorithm is effective for inventory planning but may require additional adjustments, such as incorporating seasonal trends or promotional impacts, to improve precision.

5. Discussions

The results of this study demonstrate that the Apriori algorithm effectively identifies frequent itemsets and association rules, making it a valuable tool for inventory planning at Agree Kopi. Compared to traditional inventory management methods that rely on fixed restocking schedules, the Apriori algorithm provides a data-driven approach that adapts dynamically to purchasing patterns. The validation process confirmed that 80% of the generated rules aligned with actual customer preferences, indicating a high level of accuracy in predicting product relationships.

When compared to previous studies, this research supports and extends earlier findings. Khairani (2023) applied Material Requirement Planning (MRP) for coffee inventory management, but the study primarily focused on lot-sizing optimization rather than identifying sales patterns. Meanwhile, Muharni (2023) utilized the Apriori algorithm to analyze sales transaction data, successfully identifying association rules but without integrating the results into an inventory planning system. Similarly, Nurhavizza & Hidayat (2022) found that Apriori could reveal frequent product combinations in a bakery store but focused more on Point-of-Sale (POS) system development rather than stock optimization.

This study builds upon these previous works by bridging the gap between transaction analysis and inventory planning. Unlike prior research that applied Apriori mainly for sales insights, this study integrates the algorithm into stock management, ensuring that frequently purchased items are well-stocked while minimizing unnecessary inventory accumulation. The findings highlight that the Apriori algorithm is not only useful for understanding consumer purchasing patterns but also for optimizing inventory strategies, reducing stock wastage, and improving operational efficiency. Future improvements could incorporate seasonal trends or promotional effects to further refine the model.

6. Conclusion

Based on the application of the Ari algorithm, coffee raw material inventory planning in coffee can be done more efficiently and data-based. This algorithm helps in finding frequent purchasing patterns, thus allowing management to plan the amount of raw material inventory that matches daily demand. By taking into account the combination of raw materials that are often purchased together, coffeeshops can avoid the accumulation of unused stock and reduce the risk of running out of needed raw materials. The implementation of the Apriori algorithm for inventory planning can improve operational efficiency and ensure optimal availability of raw materials every month.

Analysis of beverage sales transaction data at Agree Coffee resulted in 69 association rules. Based on the level of trust, the distribution was obtained as follows: 2 rules with a confidence level of 80%, 6 rules with a confidence level of 89%, 1 rule with a confidence level of 90%, and 60 rules with a confidence level of 100%. All of the relationships between the found itemsets showed a positive relationship, which means the items were likely to be purchased together more often than would have been expected if they had occurred independently.

These findings suggest that most of the association rules found have a high level of confidence (above 80%), with most of the rules achieving a 100% confidence level. This indicates that the relationship between the items in the transaction in Agree Coffee is very strong and consistent, providing useful insights for inventory planning and marketing strategies. By knowing the rules of association that often occur, management can optimize raw material stocks and improve customer experience by providing products that are often paired together.

References

[1] R. N. Hafiz, D. R. Wardani, F. Aisyah, A. N. Fayyaza, M. Kandripa, and A. Pradana, "Implementasi Material Requirement Planning pada PT XYZ Implementasi Material Requirement Planning pada PT XYZ,". *Jurnal Teknologi dan Sistem Informasi*,no. December, pp. 0–17, 2024.

[2] J. Aswat, M. Maswadi, and N. Kusrini, "Efficiency of Robusta Coffee Supply Chain Performance in Kubu Raya Regency, West Kalimantan Province," *Jurnal Penelitian Pertanian Terapan*, vol. 23, no. 4, pp. 538–546, 2023, doi: 10.25181/jppt.v23i4.2976.

[3] A. Halim Hasugian, M. Siddik Hasibuan, and S. Nurhaliza Sofyan, "Apriori to Analyze Sales Patterns of Building Tools and Materials," *IT Journal of Research and Development*, vol. 7, no. 2, pp. 170–180, 2022, doi: 10.25299/itjrd.2023.10544.

[4] M. S. Sandy, H. Setiawan, U. Indahyanti, F. Sains, and D. Teknologi, "Analisis Data Mining Produk Retail Menggunakan Metode Asosiasi Dengan Menerapkan Algoritma Apriori," *Jurnal Kesatria: Penerapan Sistem Informasi*, vol. 4, no. 2, pp. 384–391, 2023, [Online]. Available: https://tunasbangsa.ac.id/pkm/index.php/kesatria/article/view/174

[5] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *Jurnal Ilmiah Informatika*, vol. 9, no. 02, pp. 75–81, 2021, doi: 10.33884/jif.v9i02.3755.

[6] G. Gunadi and D. I. Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) :," *Jurnal Telematika*, vol. 4, no. 1, pp. 118–132, 2012.

[7] I. Qoniah and A. T. Priandika, "Analisis Market Basket Untuk Menentukan Asossiasi Rule Dengan Algoritma Apriori (Studi Kasus: Tb.Menara)," *Jurnal Teknologi dan Sistem Informasi*, vol. 1, no. 2, pp. 26–33, 2020, doi: 10.33365/jtsi.v1i2.368.

[8] W. Khairani and Y. Epi, "Penerapan Metode Material Requirement Planning (MRP) Dalam Perencanaan Persediaan Kopi," *TIN Terap. Inform. Nusant.*, vol. 4, no. 7, pp. 464–473, 2023, doi: 10.47065/tin.v4i7.4676.

[9] S. Muharni and S. Andriyanto, "Penentuan Pola Penjualan Menggunakan Algoritma Apriori," *Digit. Transform. Technol.*, vol. 4, no. 1, pp. 60–71, 2024, doi: 10.47709/digitech.v4i1.3679.

[10] S. U. Nurhavizza and S. Hidayat, "Analisi Stok Pada Data Penjualan Toko Kue Menggunakan Algoritma Apriori," 2022.

[11] A. P. J. V. S. Rasid Siddik, "2024 apriori 3109-Article Text-22138-1-10-20240416," vol. 7, pp. 349–354, 2024.

[12] D. Pratiwi and J. S. Wibowo, "Implementasi Algoritme Apriori Pada Sistem Persediaan Obat Apotik Puskesmas," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 12, no. 1, pp. 214–219, 2023.

[13] H. R. S. Harahap, "Konsep Dan Fungsi Algoritma Pemrograman," *Konstanta: Jurnal Matematika dan Ilmu Pengetahuan Alam*, vol. 1, no. 3, pp. 245–257, 2023.

[14] A. Suharto, "Fundamental Bahasa Pemrograman Python," *Eureka Media Aksara*, pp. 1–25, 2023.

[15] B. Amal *et al.*, "Penilaian Mahasiswa PBSI Unsika Terhadap Keefektifan Penggunaan Google Colab Dalam Pembelajaran Coding,". *Jurnal Pendidikan Teknologi Informasi*, vol. 6, no. 1, pp. 66–76, 2025.