

## **A Design of cross-selling products based on frequent itemset mining for coffee shop business**

**Ari Hasudungan Pratama Pasaribu, M Zaky Hadi\*, Fatin Saffanah Didin, Lina Aulia**

\* Department of Industrial Engineering, Faculty of Industrial Technology, Institut Teknologi Sumatera, Indonesia, 35365. Jl. Terusan Ryacudu, Way Huwi, Jati Agung, Lampung Selatan, Lampung, Indonesia

\*✉ [zaky.hadi@ti.itera.ac.id](mailto:zaky.hadi@ti.itera.ac.id)

Submitted: 07/02/2024

Revised: 01/04/2024

Accepted: 07/04/2024

### **ABSTRACT**

Using a case study at the XYZ coffee shop in Bandar Lampung, Indonesia, this work conducts knowledge extraction on a sales dataset to create a cross-selling model for product bundling advice. Understanding Knowledge extraction was done using a frequent itemset mining technique based on an Apriori algorithm to extract a set of association rules between products. The study implemented a five-stage frequent itemset structure that encompasses business comprehension, data preparation, data exploration, model creation using the Apriori algorithm, and rules evaluation. The framework that comes from this research provides a set of bundling association rules across items for cross-selling strategies that involve many products and complex sales timeframes. Additionally, we recommended sales activities based on loyalty cards to enhance our dataset with consumer attributes and purchasing trends. Based on customized services and tailored offers based on consumers' past purchases and spending patterns, the loyalty card recommendation was created. Consequently, we may target the appropriate clients and items with our marketing initiatives.

**Keywords:** Algorithms apriori; frequent itemset mining; cross-selling; association rules; data analytics

### **1. INTRODUCTION**

Cross-selling is a marketing tactic in which vendors present customers with complementary products or services while they are completing a transaction [1]. The cross-selling technique entails providing clients with an extra product or service in exchange for their purchase, or grouping things together at a discounted price [2]. The key benefits of cross-selling include higher sales revenue, improved customer satisfaction, and, for B2B companies, higher Customer Lifetime Value (CLV) through deeper integration in a customer's business, customer-specific offers, loyalty and engagement, support for promotional programs, lower selling costs, and increased product referrals [3].

Cross-selling can be done successfully in a number of ways. The most widely used conventional method involves the salesman interpreting and analyzing the needs of the customer [4]. Although traditional cross-selling strategies have been used for a while, they have drawbacks [5]. These strategies' primary drawback is how much they depend on the salesperson's aptitude for interpreting and interpreting the wants of the consumer. Although competent salespeople can be highly successful at this, there are numerous circumstances in which they might not be able to fully understand the wants of the consumer. Sales may suffer as a result, as well as chances lost. Traditional cross-selling strategies also have the drawback of being time-consuming [6]. Salespeople often have to spend a great deal of time understanding the needs of the customer before they can even begin to recommend additional products or services [7]. Resources may be depleted by this, particularly if the salesperson is juggling several



clients at once. Additionally, customer attitudes and habits are rapidly evolving. In the big data era, cross-selling strategies driven by artificial intelligence (AI) are becoming more and more common as a means of overcoming these limitations [8]. AI-driven cross-selling strategies that leverage analytics, data mining, and machine learning are transforming the sales environment [9]. These methods provide new opportunities and benefits for all kinds of enterprises, and they are flexible and adaptable enough to deal with the ever-changing needs of their clientele. One popular method of using AI to enhance cross-selling tactics is to mine item sets to derive rules about relationships between two or more products [10].

Frequent itemset mining is the process of locating frequent item sets with the aim of establishing association rules between items in the dataset [11][12]. To extract it, market basket data is analyzed for frequently recurring if/then patterns. Then, the most significant links are identified by using criterion support and confidence (a technique known as association rules mining). A set of rules describing the frequency and strength of the relationships between the variables or items is the result of association rule mining. This will make it easier and more intuitive for the marketing manager to recognize the patterns and linkages, pinpoint the primary themes or product bundling categories, and simplify the process of making cross-selling selections.

With a case study of the marketing process at Coffee Shop XYZ in the city of Bandar Lampung, this research offers a product cross-selling strategy through the use of frequent itemset mining and association rules mining to extract knowledge. Through the use of convincing advertising campaigns and posters affixed to the storefront, the management of Coffee Shop XYZ introduced and promoted their products. This technique was no longer effective as the market and business rivalry increased because it only improved sales by 1-5% and increased the likelihood of a significant loss of 5-10% in the previous 6 months. To create a new product offering plan, management had to use past transaction datasets from the sales information system database and use them to create cross-selling promotions or product bundling strategies [13]. To offer product bundling and personalized recommendations for customers by selling complementary products or services to customers when they are making a purchase, Coffee Shop XYZ must design a cross-selling strategy through the extraction of historical transaction datasets. By doing this, the business can increase sales while providing excellent customer service.

In this study, we used a knowledge extraction and data mining approach to create a cross-selling strategy to address such issues. Two separate research goals were identified from the frequent itemset mining and association rules mining for the extraction of a cross-selling strategy. First, we extracted knowledge as association rules to find rules for product pairs that frequently (often appear together) from customer buying patterns based on historical transaction dataset. Second, as part of a cross-selling strategy, we created product suggestions based on criteria discovered through knowledge extraction for product offerings. To create a more representative rule model based on past customer purchase trends, these two goals must be met. Only then will the representative rule model be suitable for use as decision support for cross-selling tactics. This study's primary contribution is new data and a design framework for frequent itemset mining marketing techniques in coffee shop.

The remainder of this study is organized as follows: in order to accomplish our research goals, we first present our research framework. The techniques created a customized framework for the data mining project to match the stated issue and improve the viability of the suggested solution. The results of our data analytics work on the transactional dataset from Coffee Shop XYZ are presented in the next section. This includes rule extraction from the dataset and both univariate and bivariate data analysis. Additionally, we offer rule performance to help choose appropriate rules for Coffee Shop XYZ marketing initiatives. The viability of the extracted cross-selling approach employing association rules and frequent itemset mining was then tested through implementation, verification, and validation [14]. The final part, "Conclusion," summarizes the results in light of the study's goals and possible future directions.

## 2. METHOD

We created a conceptual framework for this study based on data mining and knowledge extraction techniques, which we then adjusted in the modeling portion to meet our goals. The conceptual

framework is shown in [Figure 1](#).

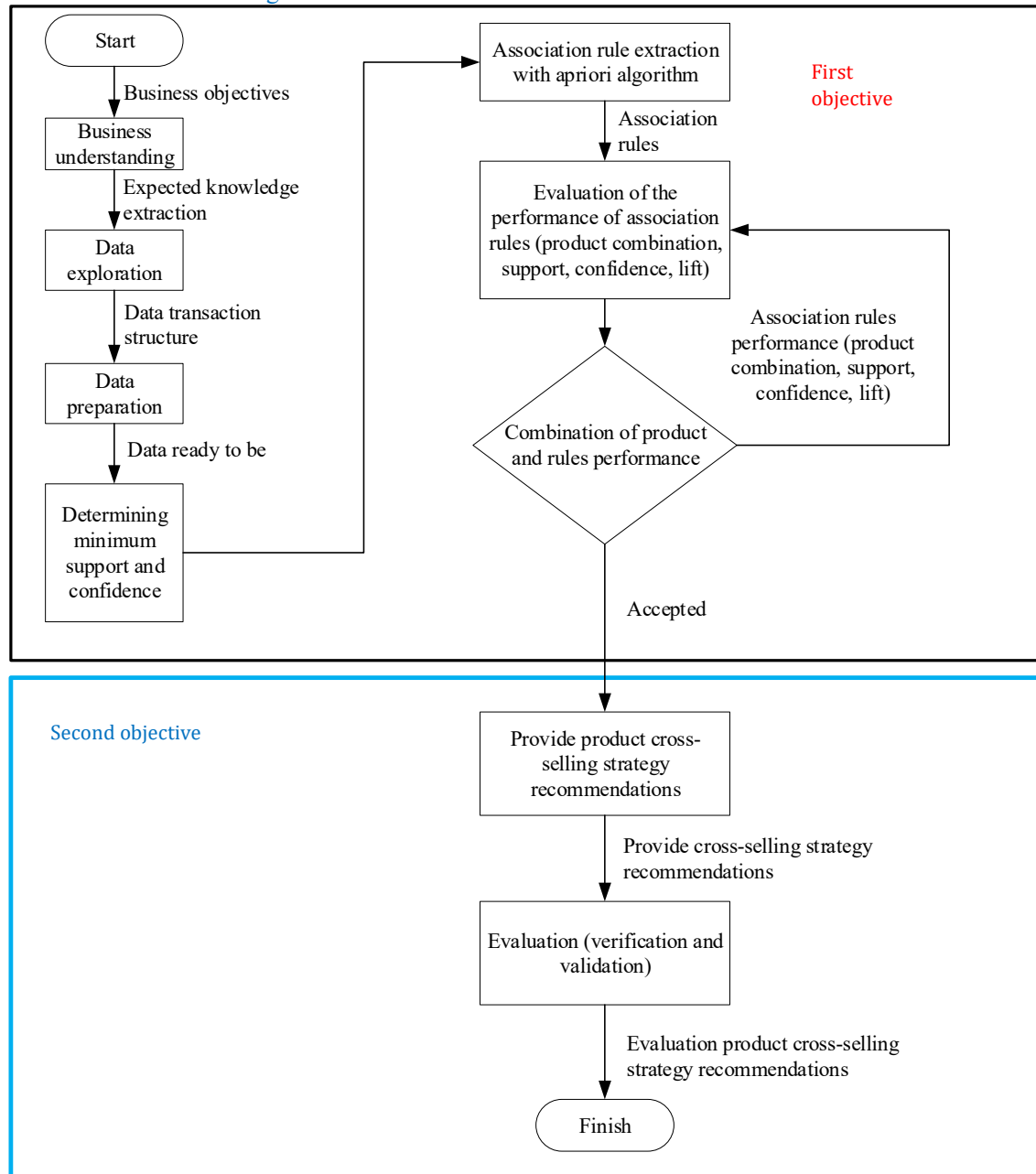


Figure 1. Research framework.

Our framework for the study is shown in [Figure 1](#). To accomplish our study aims, the framework consists of five primary stages. Understanding business processes, comprehending datasets, preparing data, modeling with the Apriori algorithm, and assessing rules are some of these activities [15].

**Data and business comprehension.** Business understanding includes setting clear goals, identifying the issue and pain spots in the business, and assessing the available resources, including time, stakeholders, and dataset requirements. The main objective is to ensure that the project is feasible to achieve the desired outcome and that it aligns with the overall business strategy and detailed business procedure. Collaboration and communication with stakeholders from many departments are also necessary at this phase to ensure that the project is in line with the larger business plan. Consequently,

the coffee shop XYZ can now more easily undertake the data analytics project. It fortifies the foundation of business knowledge and focuses attention on data research that could help accomplish the project's goals. Data mining's initial phase, known as data exploration, involves describing the dataset's attributes—like its size, quantity, and kind—using statistical techniques and data visualization to enhance our understanding of the data's nature. This phase consists of the following four tasks: Compile initial data, describe data, examine data, and ensure data quality.

Data processing. As a general guideline, 80% of the project is dedicated to data preparation. This procedure, which is also referred to as "data munging," gets the final data set or sets ready for modeling. Data selection, data cleaning, data building, data integration, and data formatting are the five jobs that make up this process. The process of choosing data specifies which datasets will be used and explains the inclusion/exclusion reasoning. Data cleaning, which includes fixing, imputing, or eliminating inaccurate values such as missing numbers and errors, is a common procedure in data mining. Mining association rules can be done with the additional attributes produced by data construction or dimensionality reduction [16]. The practice of combining data from several sources to create a new dataset is known as data integration. Reformatting data as necessary to facilitate mathematical operations is known as formatting.

The concept of data mining is based on modeling and evaluating rules, which are virtual structures that represent data grouped for predictive analysis in cross-selling strategy [17]. This phase is built using an algorithm on the preprocessed data. The transaction dataset was divided into weekend and weekday categories before being further categorized into lunchtime, afternoon, and evening. Next, using descriptive analytics, the categorized dataset was visualized. During this stage, the Apriori algorithm was applied by calculating the minimal support and confidence through a process of trial and error until a product combination was produced. Next, using the following formulation, we assessed the product combination by determining the rule's performance in terms of lift, confidence, and support.

$$Sup(X \rightarrow Y) = Sup(XY) = |t(XY)| \quad (1)$$

$$Conf(X \rightarrow Y) = P(Y|X) = \frac{P(XY)}{P(X)} = \frac{rsup(XY)}{rsup(X)} = \frac{supsup(XY)}{supsup(X)} \quad (2)$$

$$Lift(X \rightarrow Y) = \frac{P(XY)}{P(X).P(Y)} = \frac{rsup(XY)}{rsup(X).rsup(Y)} = \frac{Conf(X \rightarrow Y)}{rsup(Y)} \quad (3)$$

An association rule is represented by the equation  $X \rightarrow Y$ , where X and Y are item sets. X and Y stand for the rule's antecedent and consequent, respectively. Association rule mining, when combined with frequent item set mining, reveals interesting connections and correlations between large sets of data items between X and Y. Support in formula (1) is the number of transactions that contain both X and Y. Confidence (2) is the conditional likelihood that a transaction contains the consequent Y given that it contains the antecedent X. Lift is defined as the ratio of the joint probability of X and Y that is seen to the joint probability that would be expected if they were statistically independent. A lift of 1 indicates independence between X and Y in the event of an association rule  $X \rightarrow Y$ . If the lift value is more than 1, then X and Y have a positive association. If the lift is less than 1, there is a negative association between X and Y. The R Studio software is used to compute and extract rules.

Recommendations for a cross-selling plan, validation, and verification. Following the modeling and assessment stages, cross-selling tactics and product bundling were designed using the X-Y rules, which were then offered at a discount. Evaluations of lift ratio, confidence, and support were used to verify. The cross-selling strategy for the marketing offering program was implemented at coffee shop XYZ as a means of validation. Subsequently, the cross-selling program's evaluation was carried out based on incremental selling for every product combination. This study offers confirmation.

### 3. RESULTS AND DISCUSSION

To determine the business process in the coffee shop XYZ and to do an exploration task using the transaction dataset, business understanding, and data exploration were carried out. In contrast to data understanding, which includes data structure analysis, exploratory data analysis, univariate data analysis, and bivariate data analysis, marketing business process and product information stems from this stage.

Details on the item set and business processes. The marketing business process consists of four steps: planning a marketing strategy, implementing it at the operational level in the coffee shop, conducting a management meeting with top-level management to evaluate marketing, and recording marketing documentation in the database management system using cashier software. The main goal of this approach is to pique customers' interest in purchasing as many goods as they can. The owner of the coffee shop, the marketing manager, the cashier, and the customer at XYZ Coffee Shop participate in this process. The cashier application support is used to retrieve information from the cashier software database management system. Transaction numbers, the transaction date, the itemset, and the total price are all included in the dataset. The transaction dataset matrix is presented in Table 1 below.

**Table 1.** The result of exporting transaction dataset from the cashier software in coffee-shop XYZ.

Transaction Number	Date and Time	Itemset	Total Price (IDR)	
1	26/08/2021 11:24	french fries	15000	Same customer
2	26/08/2021 11:40	pandan milk coffee	17000	
3	26/08/2021 11:54	Milk Coffee	13000	
3	26/08/2021 11:54	Milk Coffee	13000	
3	26/08/2021 11:54	Toast	18000	
4	26/08/2021 12:30	chocolate cold	12000	
5	26/08/2021 13:07	beef black paper	25000	
.....	.....	.....	.....	

Analysis of data structures. The data structure needs to be separated into variables so that the Apriori algorithm can calculate them, as seen in Table 1. Since they weren't utilized in the extraction procedure of the rule, the transaction number in the first column and the total price in the fourth column need to be eliminated. In the meantime, midday, evening, and overnight will be distinguished in the Transaction Time column. Transactions per customer will be created from the itemset column. This will aggregate many rows of transactions that are recognized as being completed by a single customer, such as rows 3, 4, and 5, which include transactions from customers with ID 3.

**Table 2.** The data exploration result using rstudio.

Tibble [1,169 x 5] (S3:tibble/tbl_df/tbl/data.frame)				
Variables	Data Type	Value		
\$ Transaction	Num [1:1169]	1	2	3
\$ Time	POSIXct [ 1:1169]	3	3	4
\$ Item	Chr [ 1:1169]	5	6	6
\$ Classification	Chr [ 1:1169]	6	6	6
\$ Day	Chr [ 1:1169]	.....		

*Format: "2021-08-26 11:24:00" "2021-08-26 11:40:00" "2021-08-26 11:54:00" "2021-08-26 11:54:00" ...*  
*"French fries" "Pandan Milk Coffee" "Milk Coffee" "Milk Coffee" ....*  
*"Midday" "Midday" "Midday" "Midday"*  
*"Weekday" "Weekday" "Weekday" "Weekday" .....*

Analysis of exploratory data. Before performing data cleaning and computations, the data exploration analysis gathers a sizable amount of unstructured data in order to determine current values. In order to identify relationships between the objects that need software support, data exploration was done. RStudio was the software we used to process the data. Based on the number of observations and variables, as well as the types and names of the variables, this method generates the dataframe's structure. The structure of the dataset is displayed in Table 2, which has 1,169 rows of data with 5 variables and POSIXct, chr, and numeric data types for each column.

Analysis of Univariate Data. One variable is used in univariate data analysis and visualization. Using time series and categorization, this descriptive data visualization consists of two features: it shows the frequency and peak hours when customers visit the coffee shop. Figure 2 shows data classified by midday, evening, and overnight (left), as well as by date (right).

Figure 2 demonstrates that although the distribution of each date varies and tends to be uniform, customers generally purchase more in the evening and less at night. Consumers tend to visit more

frequently on weekdays than on weekends, if weekend and weekday classifications are used (Figure 3). This makes sense because office workers and students are the coffee shop's target market.

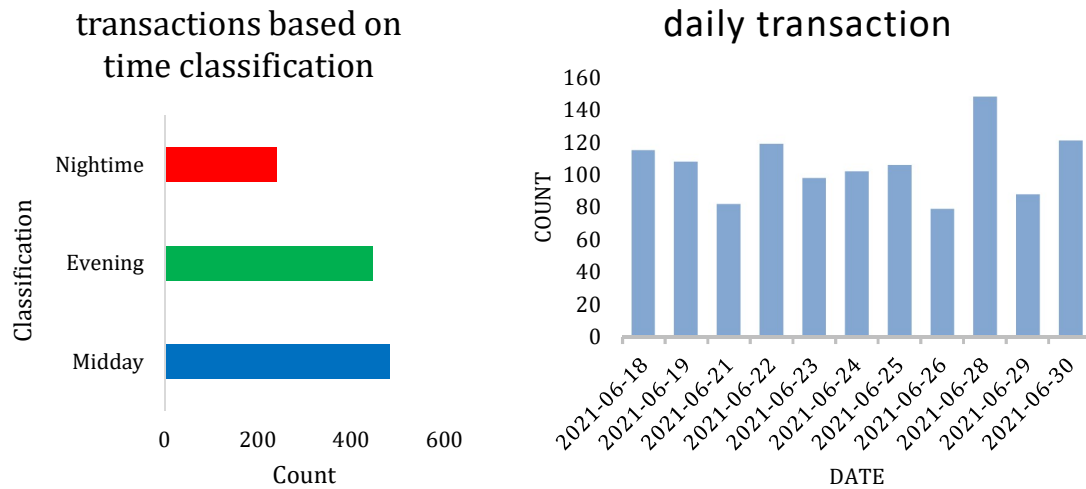


Figure 2. Descriptive data visualization based on univariate data analysis.

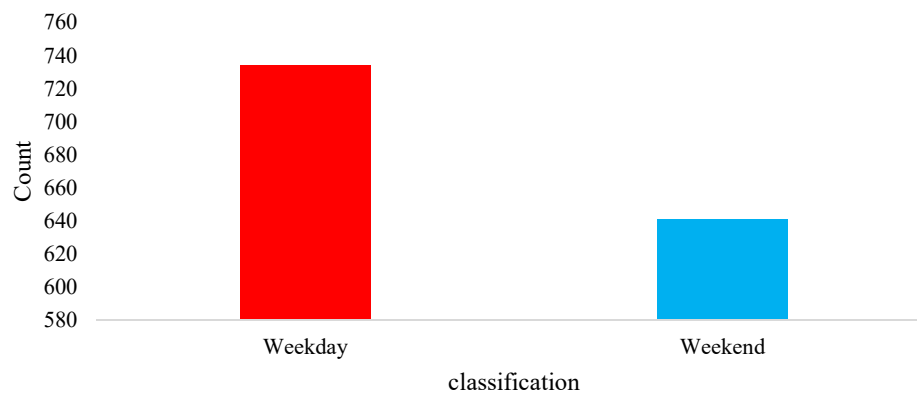


Figure 3. Distribution of transactions by day.

Analysis of bivariate data. The outcome of the data exploration stage is a descriptive visualization of the two variables. The purpose of this step was to determine how sales classes were distributed across weekdays and weekends during lunchtime, evening, and night (Figure 4). Figure 4 displays regular statistics indicating that weekday afternoons are when Coffee-Shop XYZ has the biggest sales.

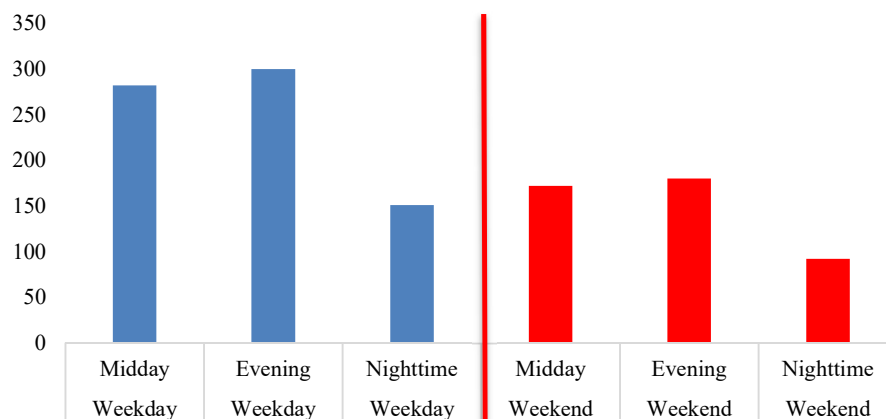


Figure 4. Bivariate analysis results for total transactions based on 2 classifications.

Data Gathering. Data preparation comes next after data exploration. This phase serves the purpose of preparing the data for processing with the Apriori algorithm. There are two categories for this stage: data transformation and data cleaning. To find noisy and incomplete data, data cleansing is done. The inspection outcome of the data cleaning procedure is shown in Table 3.

Table 3. Results of checking noisy data and incomplete data.

Table (adiksi dfsItem)				
Americano cold 38	Americano hot 6	Aqua 32	grilled chicken 7	Banana pleasure 30
Beef black paper 26	Blue curacao 10	Blue curacao 4	Caramel latte cold 36	Caramel macchiato 22
Caramel macchiato 14	Caramel macchiato cold 1	Chicken karaage 21	Chicken katsu 16	Chicken pop sambal matah 23
Chicken pop sambal matah 1	Chocolate cold 60	Churros 2	Crispy chicken skin 20	Espresso 1
French fries 54	Garlic bread 6	Hazelnut chocolate cold 66	Hazelnut latte 1	Hazelnut latte cold 66
Hazelnut latte hot 1	Hazelnut chocolate cold 1	Ice cream 1	Ice cream toast 12	Milk coffee 66
Palm milk coffee 105	Pandan milk coffee 65	Lemon tea 8	Lychee mojito 17	Lychee tea 11
Mango tea 2	Matcha green tea cold 6	Matcha green tea cold 22	Matcha green tea hot 1	Fried noodle 5
Mochaccino cold 2	Mochaccino cold 10	Mochaccino hot 2	Machos 12	Naga lychee 1
Naga Lychee 3	Naga banana 2	Onion ring 2	Oranio boba 48	Passion fruit 28
Toast 13	Canai bread 8	Sosis barbeque 4	Spaghetti aglio olio 13	Spaghetti bolognese 25
Strawberry lychee 4	Strawberry tea 5	Susu caramel 13	Hazelnut milk 14	Milk coffee with jelly 16
Milk jelly cold 2	Lychee milk 1	Susu lychee 4	Rum milk 6	Vanilla milk 2
Vanilla milk 7	Taro 21	Thai tea 1	Vanilla latte cold 6	Vanilla latte cold 31
Waffle 27				

Table 3's findings demonstrate that no noisy data or missing values were discovered, indicating that no cleaning procedure was used. Not every variable is used in the association rule mining process' computations. In order to convert the raw data from the sales database system from Table 1 to Table 4, a transaction dataset containing consumer IDs and the products they buy in each transaction is needed for this extraction process.

Table 4. Dataset transformation from table 1 into standard transaction data format.

Transaction Number	Itemset
1	French fries
2	Pandan milk coffee
3	Milk coffee, milk coffee, toast
4	Chocolate cold
5	Beef Black paper
.....	.....
560	Beef Black paper

Modeling is the next step. Using an a priori algorithm, the modeling stage explains the data processing and processing outputs. Transaction data that has undergone feature selection, data cleansing, and data input is the data that is used. The data was additionally split into weekend and weekday categories during the modeling step, and the weekend and weekday categories were further subdivided based on midday, afternoon, and night. Every computation has a visual representation.



Every Transaction. Rules comprise the "all transaction" data frame, which establishes minimum support and confidence. Following modeling, support, confidence, top item frequency, and statistical visualization computation results were acquired. Following a period of trial and error, the minimal support and confidence were found to be 0.01 and 0.2, respectively. The outcomes of mining frequent item sets and association rules for every transaction are shown in Figure 5.

The top item frequency plot, shown in Figure 5 and Table 5, displays the ten most commonly ordered items across the whole transaction data frame. Under the current circumstances, this data can be used for product sales and business concerns. Historical data indicates that coffee made from palm milk is a best-selling (frequent item). The product pair rules are also displayed in Figure 5 and Table 5. For example, the first rule states that if a customer purchases crispy chicken skin, they should also purchase hazelnut chocolate cold, with a lift of 2.8, a support of 1%, and a confidence of 30%. This suggests a strong relationship between the two items in the dataset.

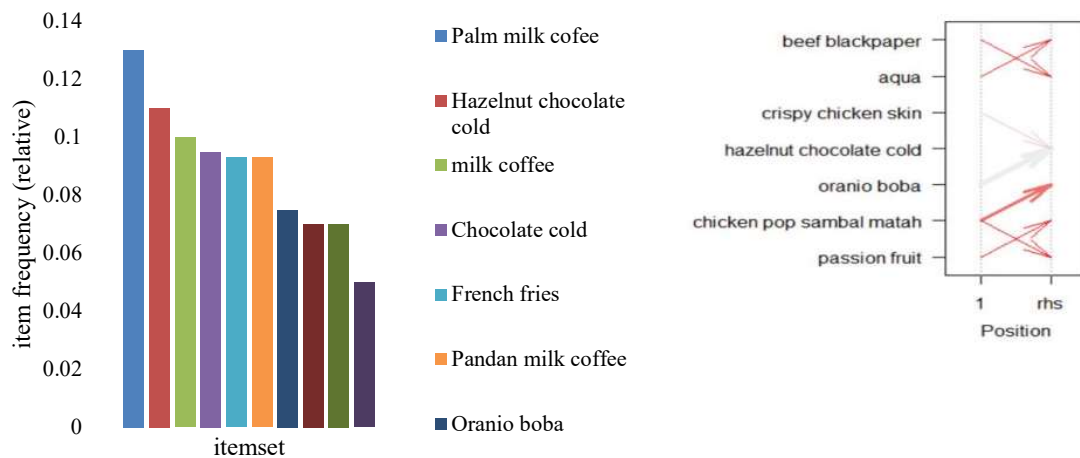


Figure 5. Results of extraction of association rules for all items.

Table 5. Matrix of rules extraction and its performance for all items.

> Inspect(rules)								
No	lhs		rhs	support	confidence	coverage	lift	count
[1]	{crispy chicken skin}	=>	{hazelnut chocolate cold}	0.0107	0.3000	0.0357	2.8050	6
[2]	{chicken pop sambal matah}	=>	{Passion fruit}	0.0107	0.2609	0.0410	5.6288	6
[3]	{passion fruit}	=>	{chicken pop sambal matah}	0.0107	0.2308	0.0464	5.6288	6
[4]	{chicken pop sambal matah}	=>	{oranio boba}	0.0125	0.3044	0.0410	4.2685	7
[5]	{beef black paper}	=>	{aqua}	0.0107	0.2727	0.0392	5.4643	6
[6]	{aqua}	=>	{beef black paper}	0.0107	0.2143	0.0499	5.4643	6
[7]	{oranio boba}	=>	{hazelnut chocolate cold}	0.0143	0.2000	0.0713	1.8700	8

Both during the week and on the weekend. The pair extraction results for weekdays are shown in Table 6, where there are seven total resulting combinations, but the weekend results show four product combination rules. A product bundle that is offered on weekday evenings is the Chicken Pop Sambal Matah bundle with passion fruit beverages, as shown in Table 7. Meanwhile, the spaghetti aglio olio bundle with passion fruit serves as an example for the weekend.



**Table 6.** Matrix of rules extraction and its performance for weekdays.

> Inspect(rules_weekdays)								
No	lhs		Rhs	support	confidence	coverage	lift	count
1	{chicken pop sambal matah}	=>	{passion fruit}	0.0220	0.6667	0.310	11.0303	4
2	{chicken katsu}	=>	{beef black paper}	0.0220	0.6667	0.3100	11.0303	4
3	{passion fruit}	=>	{chicken pop sambal matah}	0.0220	0.3636	0.0604	11.0303	4
4	{beef black paper}	=>	{chicken katsu}	0.0220	0.3636	0.0604	11.0303	4
5	{beef black paper}	=>	{aqua}	0.0220	0.3636	0.0604	5.5152	4
6	{aqua}	=>	{beef black paper}	0.0220	0.3333	0.0659	5.5152	4
7	{aqua}	=>	{hazelnut chocolate cold}	0.0220	0.3333	0.0659	3.1930	4

**Table 7.** Extraction results for weekday evening.

> Inspect(rules_weekday_evening)								
No	lhs		Rhs	support	confidence	coverage	lift	count
1	{Spaghetti aglio olio}	=>	{Passion fruit}	0.0123	0.6667	0.0184	12.0741	2
2	{Grilled chicken}	=>	{Pandan milk coffee}	0.0123	0.6667	0.0184	15.5238	2
3	{crispy chicken skin}	=>	{hazelnut chocolate cold}	0.0245	0.8000	0.0306	8.1500	4

In the meanwhile, [Table 8](#), [Table 9](#), [Table 10](#), [Table 11](#), [Table 12](#), and [Table 13](#) display the findings for the weekend and weekday lunchtime, evening, and nighttime. When a customer purchases Chicken Pop Sambal Matah, for instance, they are typically presented with Passion fruit (along with the bundling). This is an example of an offer transaction for a Saturday afternoon.

**Table 8.** Extraction results for the weekend at midday

> Inspect(rules_weekend_midday)								
No	lhs		Rhs	support	confidence	coverage	lift	count
1	{chicken pop sambal matah}	=>	{passion fruit}	0.0323	1	0.0323	20.6667	2
2	{chicken katsu}	=>	{hazelnut latte cold}	0.0323	1	0.0323	20.6667	2
3	{aqua, oranio boba}	=>	{hazelnut chocolate cold}	0.0323	1	0.0323	8.8571	2
4	{aqua, hazelnut chocolate cold}	=>	{orianio boba}	0.0323	1	0.0323	7.7500	2

> Inspect(rules_weekend_midday)								
No	lhs		Rhs	support	confidence	coverage	lift	count
5	{beef black paper, ice cream toast}	=>	{French fries}	0.0323	1	0.0323	7.7500	2
6	{French fries, ice cream toast}	=>	{beef black paper}	0.0323	1	0.0323	12.4000	2
7	{beef black paper, French fries}	=>	{ice cream toast}	0.0323	1	0.0323	20.6667	2

Table 9. Extraction results for the weekend in the evening.

> Inspect(rules_weekend_evening)								
No	lhs		Rhs	support	confidence	coverage	lift	count
1	{chicken katsu}	=>	{beef black paper}	0.0256	1	0.0256	26.00	2
2	{chicken katsu}	=>	{palm milk coffee}	0.0256	1	0.0256	6.00	2
3	{spaghetti aglio olio}	=>	{aqua}	0.0256	1	0.0256	9.75	2
4	{beef black paper, chicken katsu}	=>	{palm milk coffee}	0.0256	1	0.0256	6.00	2
5	{chicken katsu, palm milk coffee}	=>	{beef black paper}	0.0256	1	0.0256	26.00	2
6	{beef black paper, palm milk coffee}	=>	{chicken katsu}	0.0256	1	0.0256	39.00	2

Table 10. Extraction results for the weekend at night time.

> Inspect(rules_weekend_nighttime)								
No	lhs		rhs	support	confidence	coverage	lift	count
1	{nachos}	=>	{milk coffee}	0.04545	1.0000	0.0455	14.6667	2
2	{milk coffee}	=>	{nachos}	0.04545	0.6667	0.0682	14.6667	2
3	{waffle}	=>	{oranio boba}	0.04545	0.6667	0.0682	7.3333	2

Table 11. Bundle offering during transactions for weekdays at midday.

> Inspect(rules_weekday_midday)								
No	lhs		rhs	support	confidence	coverage	lift	count
1	{caramel milk}	=>	{taro}	0.0134	1	0.0134	18.6250	2
2	{beef black paper, lychee mojito}	=>	{chicken karage}	0.0134	1	0.0134	21.2857	2
3	{beef black paper, chicken karage}	=>	{lychee mojito}	0.0134	1	0.0134	29.8000	2

> Inspect(rules_weekday_midday)								
No	lhs		rhs	support	confidence	coverage	lift	count
4	{chicken karage, lychee mojito}	=>	{beef black paper}	0.0134	1	0.0134	29.8000	2

Table 12. Bundle offering during transactions for weekdays in the evening.

> Inspect (rules_weekday_nighttime)								
No	lhs		rhs	support	confidence	coverage	lift	count
1	{chicken pop sambal matah}	=>	{oranio boba}	0.0423	1.0000	0.0423	14.2000	3
2	{oranio boba}	=>	{chicken pop sambal matah}	0.0423	0.6000	0.0704	14.2000	3
3	{hazelnut latte cold}	=>	{vanilla latte cold}	0.0282	0.6667	0.0423	15.7778	2
4	{vanilla latte cold}	=>	{hazelnut latte cold}	0.0282	0.6667	0.0423	15.7778	2
5	{hazelnut latte cold}	=>	{caramel macchiato}	0.0282	0.6667	0.0423	11.8333	2
6	{vanilla latte cold}	=>	{caramel latte cold}	0.0282	0.6667	0.0423	6.7619	2

Table 13. Bundle offering during transactions for weekdays in the evening.

> Inspect (rules_weekday_evening)								
No	lhs		rhs	support	confidence	coverage	lift	count
1	{spaghetti aglio olio}	=>	{passion fruit}	0.0123	0.6667	0.0184	12.0741	2
2	{grilled chicken}	=>	{pandan milk coffee}	0.0123	0.6667	0.0184	15.5238	2
3	{crispy chicken skin}	=>	{hazelnut chocolate cold}	0.0245	0.8000	0.0307	8.1500	4

Validation, verification, and implementation. These guidelines are applied when creating bundles that provide customers with cheaper product combinations. After a week of testing, the implementation's findings showed a 30% gain in sales, which was larger than the previous weeks' sales, which had been trending downward with a drop in the range of 5–10% and an increase of 1-5% in total product sales. For the product offering process, Coffee-Shop XYZ introduced a loyalty card to make it more customized for each customer.

#### 4. CONCLUSION

The research's goals have led to the following conclusions being reached. First, products with positive connections and frequent item sets are those that exhibit time-varying rules, as determined by using knowledge extraction and association rules mining to Coffee Shop XYZ transaction data. Based on weekend and weekday periods, this study offers suggestions for lunchtime, evening, and nighttime. The design approaches for product bundling both increase sales. By offering loyalty cards, personalization must be made to facilitate the extraction of association rules. One limitation of this research is that the tailored process was not conducted based on customer classification, which determines what is lucrative and non-profitable based on the loyalty card operations that have been

completed. For this reason, cohort analysis, customer segmentation, and customer personalization are examples of advanced analytics procedures that are needed.

## REFERENCES

- [1] T. D. Astuti, T. I. Hermanto, and I. Kaniawulan, "Analisa data mining menggunakan algoritma apriori untuk meningkatkan cross selling dan up selling (Studi kasus Rumah Makan Mas Nur Purwakarta)," *J. Teknol. dan Inf.*, vol. 6, no. 2, pp. 69–77, 2016.
- [2] L. Zhang, J. Priestley, J. Demaio, S. Ni, and X. Tian, "Measuring Customer Similarity and Identifying Cross-Selling Products by Community Detection," *Big Data*, vol. 9, no. 2, pp. 132–143, 2021, doi: 10.1089/big.2020.0044.
- [3] S. Li, B. Sun, and A. L. Montgomery, *Cross-selling the right product to the right customer at the right time*, vol. 48, no. 4, 2011. doi: 10.1509/jmkr.48.4.683.
- [4] G. K. Hunter and W. D. Perreault, "Making sales technology effective," *J. Mark.*, vol. 71, no. 1, pp. 16–34, 2007, doi: 10.1509/jmkg.71.1.16.
- [5] S. Li, B. Sun, and A. L. Montgomery, "Cross-selling the right product to the right customer at the right time," *J. Mark. Res.*, vol. 48, no. 4, pp. 683–700, 2011, doi: 10.1509/jmkr.48.4.683.
- [6] İ. E. Kalkan and C. Şahin, "Evaluating cross-selling opportunities with recurrent neural networks on retail marketing," *Neural Comput. Appl.*, vol. 35, no. 8, pp. 6247–6263, 2023, doi: 10.1007/s00521-022-08019-1.
- [7] J. Ben Schafer, J. A. Konstan, and J. Riedl, "E-commerce recommendation applications," *Data Min. Knowl. Discov.*, vol. 5, no. 1–2, pp. 115–153, 2001, doi: 10.1007/978-1-4615-1627-9\_6.
- [8] M. Moradi and M. Dass, "Applications of artificial intelligence in B2B marketing: Challenges and future directions," *Ind. Mark. Manag.*, vol. 107, no. September 2020, pp. 300–314, 2022, doi: 10.1016/j.indmarman.2022.10.016.
- [9] J. Bharadiya and J. P. Bharadiya, "Machine Learning and AI in Business Intelligence: Trends and Opportunities," *Int. J. Comput.*, vol. 48, no. 1, pp. 123–134, 2023, [Online]. Available: <https://www.researchgate.net/publication/371902170>
- [10] V. Kannan, D. Klötzl, and A. Sharma, "Visualizing the Interactions and Relationships from Sales Data and Data-Driven Automatic Product Bundling to Increase Cross-Selling," 2019.
- [11] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal, "Efficient mining of association rules using closed itemset lattices," *Inf. Syst.*, vol. 24, no. 1, pp. 25–46, 1999, doi: 10.1016/S0306-4379(99)00003-4.
- [12] M. Hahsler, B. Grün, and K. Hornik, "Arules - A computational environment for mining association rules and frequent item sets," *J. Stat. Softw.*, vol. 14, no. 2005, pp. 1–25, 2005, doi: 10.18637/jss.v014.i15.
- [13] C. Schmitz, Y. C. Lee, and G. L. Lilien, "Cross-selling performance in complex selling contexts: An examination of supervisory- and compensation-based controls," *J. Mark.*, vol. 78, no. 3, pp. 1–19, 2014, doi: 10.1509/jm.12.0421.
- [14] T. Brijs, G. Swinnen, K. Vanhoof, and G. Wets, "Building an Association Rules Framework," *Data Min. Knowl. Discov.*, vol. 8, no. 1, pp. 7–23, 2004, [Online]. Available: <http://0-search.proquest.com.opac.sfsu.edu/docview/230135438>
- [15] W. Van Der Aalst, "Process mining: Overview and opportunities," *ACM Trans. Manag. Inf. Syst.*, vol. 3, no. 2, pp. 1–17, 2012, doi: 10.1145/2229156.2229157.
- [16] M. Usman, R. Pears, and A. C. M. Fong, "Discovering diverse association rules from multidimensional schema," *Expert Syst. Appl.*, vol. 40, no. 15, pp. 5975–5996, 2013, doi: 10.1016/j.eswa.2013.05.031.
- [17] A. Rahman and M. N. A. Khan, "An Assessment of Data Mining Based CRM Techniques for Enhancing Profitability," *Int. J. Educ. Manag. Eng.*, vol. 7, no. 2, pp. 30–40, 2017, doi: 10.5815/ijeme.2017.02.04.