Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.3, June 2025, pp. 756~764
e-ISSN: 2548-6861

Implementation of FP-Growth Algorithms for Promo Package Determination in a Scooter Motorcycle Workshop Business

I Gusti Ngurah Bagus Picessa Kresna Mandala [1]*, I Ketut Adi Purnawan [2]*, I Made Agus Dwi Suarjaya [3]*
* Teknologi Informasi, Universitas Udayana
picessakresnamandala@student.unud.ac.id [1], adipurnawan@unud.ac.id [2], agussuarjaya@it.unud.ac.id [3]

Article Info

Article history:

Received 2025-04-25
Revised 2025-05-03
Accepted 2025-05-06

Keyword:

FP-Growth Algorithm,
Bundle Package,
Association rules,
Data Mining,
Customer Satisfaction

ABSTRACT

This study applies the FP-Growth algorithm to design bundled promotions for a scooter motorcycle accessory store and workshop in Denpasar, Bali. FP-Growth was chosen for its efficiency in mining frequent itemsets without generating candidate sets. From 23,381 transaction records (January-August 2024), the algorithm identified 16 association rules using a minimum support of 1% and confidence of 50%. These rules were selected based on lift values and product relevance. One notable example is the association between "BAUT TITANIUM GR5 M10 X 60" and "BAUT TITANIUM GR5 M8X50", which had a lift of 47.814, indicating a very strong co-purchase relationship. These high-lift combinations present valuable opportunities for bundling and targeted point-of-sale offers. The algorithm performed efficiently, with a runtime of just 0.1354 seconds and 402.6 MB of memory usage. Bundles based on these associations were presented to customers, and feedback was collected through a Customer Satisfaction (CSAT) survey involving 56 recent buyers. The survey yielded a high CSAT score of 83.93%, demonstrating customer satisfaction with the bundles' relevance and appeal. These results confirm that FP-Growth can effectively inform promotional strategies by identifying strong product pairings that align with actual purchasing behavior. Strategically promoting such bundles not only enhances customer experience but also encourages multi-item purchases. This data-driven bundling approach is practical and profitable for medium-sized retail businesses, ultimately supporting the goal of increasing the Average Order Value.

I. Introduction

The rapid advancement of technology must be optimally utilized across various sectors, including the trade sector. Information technology has become an essential tool for business actors to increase operational efficiency, accelerate decision-making processes, build customer loyalty, and boost sales value[1]. One strategy that can be applied to achieve these goals is the development of effective and data-driven promotional efforts[2].

Promotion is a part of the marketing mix that functions to convey product benefits and persuade consumers to make a purchase[3]. Well-designed promotional strategies can enhance brand awareness, shape consumer preferences, and retain customer loyalty[4]. In practice, many small and medium enterprises still rely on intuition when designing their promotional strategies, without leveraging historical sales data that actually holds great potential for deeper analysis[5].

As the number of customers and transactions increases, the collected sales data becomes a valuable asset that can be used to discover consumer purchasing patterns[6]. One approach that can be used to analyze this data is data mining techniques, specifically association rule mining[7]. This technique aims to find associative rules between products that are frequently purchased together, which can then be used to design promotional packages that match customer preferences.

Bundle packages or promo packages are a marketing strategy in which several products are combined into a single unit sold at one price[8]. This strategy not only simplifies the purchasing process but also encourages an increase in the transaction's total sales volume[9]. To design effective promotional bundles, a data-driven approach is needed to identify the most frequently purchased product combinations[10].

Several algorithms are widely used in association rule mining, with Apriori, ECLAT, and FP-Growth being among the most prominent[11]. Each of these algorithms has distinct characteristics in terms of performance and scalability. Apriori, while conceptually straightforward, tends to be less efficient because it requires the generation of all candidate itemsets and multiple passes over the dataset. This can become computationally expensive as the dataset grows.

The ECLAT algorithm addresses several limitations of the Apriori algorithm by utilizing a vertical data format, which enables the algorithm to perform intersections on transaction ID lists rather than conducting multiple scans of the entire database. This approach can improve computational efficiency. However, ECLAT tends to consume a substantial amount of memory as the dataset size increases, primarily due to the need to store vertical representations of itemsets. For example, when applied to a dataset consisting of 8,416 transactions with a minimum support threshold of 0.3, ECLAT required approximately 146.81 MB of memory, whereas Apriori consumed only 26.84 MB[11].

In contrast, the FP-Growth algorithm eliminates the candidate generation process by constructing a compact data structure called the FP-Tree, which compresses the dataset by identifying and grouping common prefixes[12][13]. This method not only accelerates the mining process but also significantly reduces memory usage by avoiding redundant storage of itemsets. Under the same dataset conditions, FP-Growth utilized only 4.21 MB of memory substantially lower than both Apriori and ECLAT indicating superior memory efficiency[11].

A performance comparison conducted on datasets with more than 8,416 transactions revealed that Apriori generated 505 frequent itemsets in 1,016 milliseconds, ECLAT in 328 milliseconds, and FP-Growth in 379 milliseconds[11]. While ECLAT exhibited the fastest execution time, the time difference between ECLAT and FP-Growth was relatively small. Given FP-Growth's considerably lower memory consumption, the algorithm offers a favorable trade-off between execution speed and resource efficiency. This makes FP-Growth particularly suitable for small and medium-sized enterprises (SMEs), which often operate under constraints in computational capacity and memory resources.

This study utilizes the FP-Growth algorithm to analyze 23,381 transaction records from a scooter motorcycle accessory store and workshop in Denpasar, Bali. The analysis focuses on identifying frequent product combinations by evaluating the algorithm's performance in terms of memory efficiency, execution time, and the quality of the association rules generated. These frequent itemsets serve as the foundation for designing promotional bundles that reflect actual customer purchasing behavior[14].

The ultimate goal of this research is to support data-driven strategies that increase the Average Order Value (AOV). By understanding which products are commonly bought together, the business can create more effective, targeted promotions that encourage customers to purchase more items per transaction[15]. The findings demonstrate the practical value of FP-Growth in helping small and medium-sized enterprises (SMEs) optimize their marketing efforts, improve customer satisfaction, and drive higher sales, even within resource-limited environments. These insights also provide a foundation for continuous improvement in promotional planning based on evolving customer behavior.

II. METHODOLOGY

This study applies the FP-Growth algorithm to sales data using a case study of a scooter motorcycle workshop and accessory store located in Denpasar, Bali. The store offers a variety of accessories and spare parts that are often purchased together by customers. Data was collected through direct interviews with the workshop owner and obtained from store transaction records. This study uses the KDD (Knowledge Discovery in Databases) methodology for the data analysis process[16], which consists of five main stages.

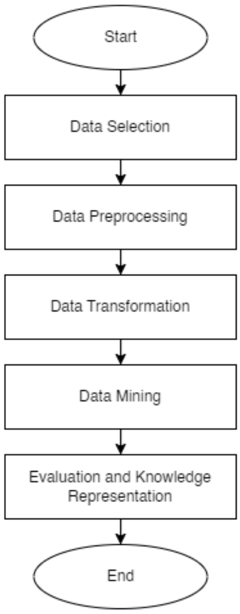

Figure 1. KDD Diagram

A. Data Selection

The data used in this study consists of eight months of sales transaction records, obtained from store documentation and direct interviews with the owner of a scooter motorcycle workshop located in Denpasar, Bali. The data includes information about the products purchased in each transaction. The dataset encompasses 23,381 transaction records, each containing a list of purchased products.

B. Data Preprocessing

Selected data is cleaned by removing duplicates, correcting typographical errors, and eliminating incomplete records. This process aims to ensure that the data used is valid, consistent, and reliable for further analysis. In this stage, outliers and anomalies in the dataset are also identified and removed to ensure that the resulting itemsets reflect true purchasing patterns.

C. Data Transformation

The cleaned data is then transformed into an itemset format per transaction to be processed by the FP-Growth algorithm. Each transaction is represented as a row containing a list of products purchased together. The data is then converted into a binary format, where each product is represented as a column with values of 1 (if the product is purchased in the transaction) or 0 (if not), to facilitate the data mining process[17].

D. Data Mining

The pattern mining process in this study is performed using the FP-Growth algorithm to identify frequent itemsets, which are combinations of products that are often purchased together by customers. The steps in the data mining phase are as follows:

1) *Calculating Minimum Support and Confidence*: The initial stage of association rule mining involves determining appropriate minimum support and confidence thresholds to filter out insignificant item combinations. These thresholds are essential for ensuring that the resulting rules are both meaningful and manageable. To address this, many studies recommend setting the minimum support below 10% when applying the FP-Growth algorithm to large datasets, as this approach effectively captures diverse and relevant item combinations without compromising computational efficiency[18]. This study follows a similar empirical strategy by testing several support and confidence threshold combinations to identify the most optimal values for the given dataset.

2) *Algorithm Performance Evaluation*: The performance of the algorithm is evaluated to measure its efficiency in processing the data. This evaluation assesses both the processing time and memory usage required for the algorithm to generate the association rules[19]. The FP-Growth algorithm is specifically chosen for its computational efficiency. To evaluate its performance, the execution time and memory usage were measured during the process of generating association rules from the transaction data.

E. Evaluation and Knowledge Representation

The next step involves evaluating the model's performance and organizing the results through knowledge representation. This ensures the findings are accurate and meaningful. The following section explains the evaluation process and representation techniques in more detail:

1) *Evaluation:* The evaluation phase is crucial for assessing the effectiveness of the FP-Growth algorithm in identifying valuable patterns. The performance of the resulting association rules is measured using key metrics such as Lift Ratio, which quantifies the strength of a rule by comparing the observed co-occurrence of items with what would be expected if they were independent, and Leverage, which measures the actual increase in probability of co-occurrence relative to independence. These metrics help identify rules that are not only statistically strong but also commercially relevant.

2) *Knowledge Representation:* After generating the frequent itemsets and association rules, the next step is to present the recommended promotional packages directly to customers through a simple and user-friendly website. This website displays combinations of products that are frequently purchased together, designed as bundled promotions. To evaluate the effectiveness and relevance of these bundles, a Customer Satisfaction Survey (CSAT) is conducted. The survey uses a 5-point Likert scale, where respondents rate aspects such as usefulness, appeal, and overall satisfaction with the bundled offers (1 = Very Dissatisfied, 5 = Very Satisfied)[20]. The CSAT score is calculated using the following equation:

$$\frac{\text{Number of Satisfied Responses}}{\text{Total Number of Responses}} \times 100 \qquad (1)$$

This equation provides a percentage that reflects the level of customer satisfaction with the promotional offers, which can be used to assess the effectiveness of the implemented strategies.

III. Result and Discussion

The results reveal the most frequently purchased combinations of items, along with the association rules generated from these combinations based on the analyzed transaction data.

A. Data Selection

The transaction data used in this study spans from January 2 to August 31, 2024, consisting of a total of 23,381 rows. These transaction records form the basis for the data mining analysis. However, because the dataset only captures eight-month timeframe, it may not comprehensively reflect customer purchasing behaviour throughout the entire year. The sales transaction data utilized in this study are presented in Table I.

TABEL I
SALES TRANSACTION DATA

Invoice	Date	Item Name	Item SKU	Iten Qtt	Item Price
#016035	02/01/ 2024	BAUT TINANIU M GR 5 M8X40	SKU: 000831	2 PCS	Rp. 60.000
#016035	02/01/ 2024	BAUT TINANIU M GR 5 M6X20	SKU: 000824	5 PCS	Rp. 35.000
#016036	02/01/ 2024	LUIGY TUTUP SPION MODEL TONY CHROME	SKU: 000932	1 PCS	Rp. 250.000
…					
#031755	31/08/ 2024	PINTIL BAN	SKU: 000419	2 PCS	Rp. 10.000

The columns used are: Invoice, Date, Item Name, Item SKU, Item Quantity, and Item Price. The original dataset is stored in CSV format, with each row corresponding to a single transaction containing individual items.

B. Data Preprocessing

The initial dataset consisted of 23,381 rows of data. A series of preprocessing steps was carried out to ensure data quality and consistency. These steps included:

1) *Removing Null Values:* Transactions with incomplete or missing data were eliminated to maintain the integrity of the dataset.
2) *Remove Duplicate:* Duplicate entries were identified and removed to avoid redundancy.
3) *Fixing Formatting Errors:* Inconsistencies such as extra spaces or misformatted inputs were standardized to ensure uniformity across the dataset.

Additionally, any outliers or erroneous data points were detected and addressed to prevent them from skewing the results. Finally, the dataset was transformed into a suitable format for analysis, ensuring all necessary attributes were properly structured for the FP-Growth algorithm.

TABEL II
CLEAN SALES TRANSACTION DATA

Invoice	Item Name	Item SKU	Item Price
#016035	BAUT TINANIUM GR 5 M8X40	000831	60.000
#016035	BAUT TINANIUM GR 5 M6X20	000824	35.000
#016036	LUIGY TUTUP SPION MODEL TONY CHROME	000932	250.000
…			
#031755	PINTIL BAN	000419	10.000

The preprocessed sales transaction data are presented in Table II. As a result of this cleaning process, the dataset was reduced to 20,521 valid records. These cleaned transactions were then transformed into receipt-based data to be used in the next stages of analysis

C. Data Transformation

The process of generating receipt-based data was carried out by grouping transactions that shared the same invoice number, indicating that the items were purchased together in a single transaction. The preprocessed transaction data used in this study are presented in Table III.

TABEL III
PREPROCESSED TRANSACTION DATA

Invoice Number	ID Product
#016035	000230,000826
#016036	000001,000466
#016037	000346,000347,000376,000768
#016038	000243,000244,000349,001012
#016039	000138

The resulting receipt-based data consists of a total of 11,671 unique invoices, with each invoice representing a complete transaction. Each item within a transaction is identified by its SKU, and this data is then stored in CSV format for further analysis. The next step involves transforming the prepared transactional data into a binary matrix format, where each row represents a transaction and each column corresponds to a specific item. The binary dataframe representing the presence or absence of items in each transaction is shown in Table IV.

TABEL IV
BINARY DATAFRAME

Invoice Number	000001	000002	…	001287	001293
#016035	0	1		0	1
#016036	0	0		0	0
#016037	0	0		0	0
#016038	0	1		1	0
#016039	0	1		0	0
…					
#031755	0	0		0	1

In this matrix, the presence of an item in a transaction is marked with a value of 1, while its absence is represented by 0. This format is essential for processing the data using the FP-Growth algorithm.

D. Data Mining

The next step is calculating the minimum support and confidence to ensure that the generated association rules meet the desired thresholds for performance and relevance. Following this, the process continues with algorithm performance evaluation, where insights are organized and presented in a form that supports practical application and informed decision-making.

1) *Calculating Minimum Support and Confidence*: Initially, the minimum support value is determined empirically by testing a range of parameter values to identify the most relevant and insightful association rules based on the observed transaction data[6]. This data-driven approach ensures that the threshold is not chosen arbitrarily but is grounded in actual patterns within the dataset. The number of association rules generated at specific support and confidence thresholds is presented in Table V.

TABEL V
NUMBER OF RULES AT A CERTAIN SUPPORT AND CONFIDENCE LEVEL

No	Support	Confidence	Assosiation Rules
1	1%	40%	19
		50%	16
		60%	10
		70%	6
		80%	5
2	2%	40%	10
		50%	10
		60%	7
		70%	5
		80%	4
3	3%	-	3
4	4%	-	3
5	5%	-	3
6	6%	-	3
7	7%	-	3
8	8%	-	3
9	9%	-	3
10	10%	-	3

Based on empirical analysis, the number of association rules generated across various combinations of support and confidence levels varies significantly, illustrating that these parameters greatly influence the quantity of rules produced. Notably, when the support threshold exceeds 3%, the number of rules remains consistent at 3, regardless of the confidence level, indicating diminishing returns in rule discovery beyond this point. The selection of the optimal parameters was grounded in empirical experimentation, followed by validation through interviews with the business owner to ensure practical relevance and alignment with real-world considerations. This methodical approach ensured that the final configuration was both analytically robust and contextually appropriate for the business environment.

The empirical analysis identified a minimum support threshold of 1% (0.01) and a minimum confidence level of 50% (0.5) as the optimal parameters, resulting in the generation of 16 association rules. The 1% support threshold was selected to capture only frequently purchased items, thereby excluding infrequent transactions and ensuring the relevance of the resulting itemsets. Additionally, the 50% confidence level was chosen to provide a reliable indication of item cooccurrence, ensuring that the generated rules reflected meaningful and actionable purchasing relationships. While a configuration of 1% support and 40% confidence resulted in a higher number of rules (19), the 50% confidence level was preferred by the business owner due to its balance between rule quantity and interpretability.

2) *Algorithm Performance Evaluation*: Initially, the minimum support value is determined empirically by testing a range of parameter values to identify the most relevant and insightful association rules based on transaction data containing 633 unique items. The runtime and memory usage of the FP-Growth algorithm at specific support and confidence thresholds are presented in Table VI.

TABEL VI
RUNTIME AND MEMORY USAGE

Unique Items	Runtime (ms)	Memory Usage (MB)	Number of Rules
633	0.1354	402.6	16

To evaluate the performance of the FP-Growth algorithm, an experiment was conducted using a dataset containing 633 unique items. With a runtime of approximately 0.1354 milliseconds and memory usage of 402.6 MB, the algorithm was able to efficiently process the data and produce 16 strong association rules. These results highlight the algorithm's capability to handle relatively large itemsets with minimal computational cost. The frequent itemsets generated using the FP-Growth algorithm with a minimum support threshold of 1% are presented in Table VII.

TABEL VII
FREQUENT ITEMSET

No	Itemset	Support
1	PAKET SERVICE MOTUL VESPA I-GET & 3V	0.28122
2	OLI GARDAN MOTUL 80W90	0.23693
3	PAKET SERVICE MOTUL VESPA I-GET & 3V, OLI GARDAN MOTUL 80W90	0.19331
..	…	…
..	…	…
..	…	…
65	MUR PLONG TITANIUM M8 GR5, BAUT TITANIUM GR5 M8X50	0.01018

Once the most frequent itemsets have been identified, the FP-Growth algorithm is used to generate association rules. These rules are then filtered using a minimum confidence threshold of 50% (0.5). Like the support value, this confidence level was selected through grid search optimization to ensure the reliability of the results, rather than being set arbitrarily. This step ensures that the resulting rules offer actionable insights that can support product bundling decisions and enhance marketing strategies. The association rules generated using the FP-Growth algorithm with a minimum support of 1% and a confidence threshold of 50% are presented in Table VIII.

TABEL VIII
ASSOCIATION RULES

No	Antecedent	Consequent	Supp	Conf
1	OLI GARDAN MOTUL 80W90	PAKET SERVICE MOTUL VESPA I-GET & 3V	0.1933	0.6874
2	PAKET SERVICE MOTUL VESPA I-GET & 3V	OLI GARDAN MOTUL 80W90	0.1933	0.8158
3	FILTER OLI ORIGINAL PIAGGIO	PAKET SERVICE MOTUL VESPA I-GET & 3V	0.0245	0.6491
…	…	...	…	…
…	…	...	…	…
…	…	...	…	…
16	BAUT TITANIUM GR5 M8X50	MUR PLONG TITANIUM M8 GR5	0.0128	0.8787

Antecedents and consequents are used to describe the associative relationship between purchased products. The 16 association rules generated are considered suitable for developing promotional bundling strategies because they meet the empirically determined minimum thresholds of 1% support and 50% confidence. These values ensure that the rules reflect purchasing patterns that are both frequent and reliable, making them not only statistically meaningful but also commercially viable. In association rule mining, antecedents and consequents represent product relationships, where the confidence value indicates the likelihood that a customer will purchase the consequent item when the antecedent is present in a transaction. This provides valuable insight into customer buying behavior and helps identify natural product groupings for bundling.

One example of a rule that meets these criteria involves the purchase of “Filter Oli Original Piaggio” as the antecedent and “Oli Gardan Motul 80W90” as the consequent. The support and confidence values of this rule confirm that a notable portion of customers who purchase the Piaggio oil filter also tend to buy the Motul gear oil. This association is logical, as both items are commonly needed together for routine scooter maintenance. Bundling them together not only simplifies the purchasing process for customers but also encourages the purchase of complementary products, potentially increasing the average order value. Moreover, such a bundle can improve perceived value, especially when offered with a small discount or promotional incentive. Overall, these 16 association rules form a data-driven foundation for creating effective promotional bundles tailored to customer behavior and service patterns.

E. Evaluation and Knowledge Representation

The next step is to evaluate the association rules that have been generated, ensuring they meet the desired thresholds for performance and relevance. Following this evaluation, the process continues by making a knowledge representation, where the insights are organized and presented in a form that supports practical application and informed decision-making.

1) *Evaluation*: Next, the following step is to evaluate the generated association rules using the lift ratio. Higher lift values indicate stronger associations between item pairs. A lift above 1 suggests a positive relationship. In this section, the top five rules with the highest lift values will be discussed.

TABEL IX
TOP 5 HIGHEST LIFT VALUES

No	Antecedent	Consequent	Lift	Lev	Conv
1	BAUT TITANIUM GR5 M10 X 60	BAUT TITANIUM GR5 M8X50	47.814	0.012	8.098
2	PER KOPLING KAWAHARA 1500RPM	PER TORSI KAWAHARA 1500RPM	31.893	0.009	6.690
3	MUR PLONG TITANIUM M8 GR5	BAUT TITANIUM GR5 M8X50	31.681	0.010	2.204
4	ROLLER KAWAHARA 11GRAM 150CC	PIAGGIO - SLIDING PAD 150CC ORIGINAL	7.885	0.016	2.925
5	ROLLER KAWAHARA 10GRAM 150CC	PIAGGIO - SLIDING PAD 150CC ORIGINAL	5.975	0.011	1.907

The top five association rules with the highest lift values, derived from the FP-Growth algorithm, reveal the strongest relationships between frequently purchased products in the dataset. These rules are critical for identifying meaningful product pairings that can be leveraged for effective promotional strategies.

The rule *"BAUT TITANIUM GR5 M10 X 60 → BAUT TITANIUM GR5 M8X50"* stands out with the highest lift value of 47.814 and leverage of 0.012, indicating a very strong association. This means customers who purchase the M10 bolt are significantly more likely to also purchase the M8 bolt. With a conviction of 8.098, this rule has strong predictive power, making it an ideal candidate for bundling such as offering the second bolt at a discount at checkout.

The second rule, *"PER KOPLING KAWAHARA 1500RPM → PER TORSI KAWAHARA 1500RPM"*, has a lift of 31.893 and leverage of 0.009, suggesting a strong complementary relationship between performance clutch springs and torsion springs, commonly used together in scooter modifications. The conviction of 6.690 reinforces its potential as a targeted bundle for performance-oriented customers.

Next, the rule *"MUR PLONG TITANIUM M8 GR5 → BAUT TITANIUM GR5 M8X50"* shows a lift of 31.681 and leverage of 0.010. This pairing makes sense mechanically and practically nuts and bolts with matching specifications are typically used together making it an intuitive upsell package that can be offered during checkout to increase average order value (AOV).

Two additional rules show associations between rollers and sliding pads. The rules *"ROLLER KAWAHARA 11GRAM 150CC → PIAGGIO SLIDING PAD 150CC ORIGINAL"* and *"ROLLER KAWAHARA 10GRAM 150CC → PIAGGIO SLIDING PAD 150CC ORIGINAL"* have lift values of 7.885 and 5.975 respectively. Their leverage values (0.016 and 0.011) and convictions (2.925 and 1.907) confirm that these components, both used in CVT systems, are often replaced together. These items can be bundled as a maintenance or upgrade kit, offered during a promotional campaign or service period.

These insights not only demonstrate statistically strong relationships but also reflect real-world usage patterns, validated through business stakeholder input. The proposed bundling strategies focus on checkout-stage recommendations (point-of-sale upselling) to increase AOV. For instance, when a customer purchases a primary item, a matching complementary item from the rules can be suggested with a time-limited discount or combo price. This approach leverages strong association metrics like lift and conviction while aligning with actual customer behavior and mechanic compatibility, ensuring both business and customer value.

2) *Knowledge Representation:* In the knowledge representation process, two key activities were conducted: the implementation of bundling strategies and the distribution of a customer satisfaction survey.

Figure 2. Bundling Website

These steps were carried out to confirm the effectiveness of the association rule-based recommendations in a real-world setting. The bundling implementation was presented through a streamlit interface that visually showcased product combinations generated by the FP-Growth algorithm. Each bundle was designed not only to reflect the most frequent itemsets but also to include discount offers, making the promotional packages more attractive to customers.

The customer satisfaction survey was conducted offline through direct interviews with 56 respondents at a scooter motorcycle accessory store and workshop in Denpasar, Bali. The data collection took place over a 12-day period between March and April, without a fixed daily schedule. Interviews were conducted for approximately one hour at the end of each store’s closing time. Respondents were customers who had recently completed a transaction during their visit and were willing to voluntarily participate in the survey. The decision to survey 56 respondents was based on practical considerations, including time constraints, availability of participants, and the goal of obtaining a sufficient yet manageable sample size to represent customer perceptions and gather initial feedback for evaluating the bundles.

Participants were selected using a simple random sampling approach from customers present at the store during the survey period. The satisfaction measurement was carried out after the introduction of bundling promotions to ensure that feedback reflected their actual experiences with the new promotional offers. The interview aimed to collect responses regarding their perceived value, satisfaction level, and purchase motivation related to the bundled products. CSAT was measured using a single-item question on a 5-point Likert scale (1 = Sangat Tidak Puas, 2 = Tidak Puas, 3 = Biasa, 4 = Puas, 5 = Sangat Puas), following standard CSAT methodology. The survey respons are presented in Table VI.

TABEL X
SURVEY RESPONS

Score (Likert 1–5)	Frequency	Percentage
5 (Sangat Puas)	33	58.9%
4 (Puas)	14	25%
3 (Biasa)	7	12.5%
2 (Tidak Puas)	2	3.6%
1 (Sangat Tidak Puas)	0	0%

Total of 56 respondents participated in the customer satisfaction survey regarding the promotional bundling offers. Using a 5-point Likert scale (1 = Sangat Tidak Puas, 5 = Sangat Puas), the majority of respondents expressed high levels of satisfaction. Specifically, 33 respondents (58.9%) gave a score of 5 (Sangat Puas), and 14 respondents (25%) gave a score of 4 (Puas). A moderate level of satisfaction (score 3) was indicated by 7 respondents (12.5%), while only 2 respondents (3.6%) gave a score of 2 (Tidak Puas), and none gave a score of 1 (Sangat Tidak Puas). The descriptive statistics of the responses are summarized as follows:

TABEL XI
DESCRIPTIVE STATISTICS

Statistic	Value
Mean (Average Score)	4.375
Median	5
Mode	5
Standard Deviation	±0.80
Total Respondents	56

The mean satisfaction score of 4.375 suggests a high overall satisfaction level with the promotional bundles. Both the median and mode being 5 indicates that the most frequent response was the highest possible satisfaction level. Furthermore, a standard deviation of ±0.80 reflects relatively low variability in the responses, suggesting consistency in how customers perceived the promotional offers. Afterward, the CSAT Score calculation will be performed to evaluate overall customer satisfaction[14].

TABEL XII
CSAT INTERPRETATION TABLE

CSAT Score	Meaning	Warning Level
0-40%	Poor CSAT	High
40-60%	Okay CSAT	Slightly High
60-80%	Good CSAT	Low
> 80%	Excellent CSAT	Very Low

The table above contains the index score or reference table used to conclude customer satisfaction based on the CSAT score obtained. The Customer Satisfaction Score (CSAT) is calculated to determine how many respondents are satisfied with the promotional bundles offered[14]. In this case, satisfaction is defined by the number of respondents who gave a score of 4 (Puas) or 5 (Sangat Puas) on the Likert scale.

$$\frac{33+14}{56} \times 100 = 0.8461 \times 100 = 83.93\%$$

This score of 83.93% falls into the "Excellent CSAT" category, indicating that the majority of customers were highly satisfied with the promotional bundles. The results suggest that the product bundling approach was well-received, with many customers expressing appreciation for the relevance and appeal of the item combinations. Moreover, the bundling strategy is practically implementable in the store's promotional planning. In addition to increasing customer satisfaction, the bundling is expected to positively influence Average Order Value (AOV) by encouraging customers to purchase more items per transaction. This indicates that data-driven bundling can be both an effective marketing tool and a profitable strategy for increasing sales performance.

IV. CONCLUSION

This study demonstrates the effective application of the FP-Growth algorithm in designing bundle package promotions for a scooter motorcycle accessory store and workshop in Denpasar, Bali. By analyzing 23,381 transaction records from January to August 2024, the algorithm uncovered statistically significant associations among frequently co-purchased items. Notable examples include the strong pairing of "BAUT TITANIUM GR5 M10 X 60" and "BAUT TITANIUM GR5 M8X50" with a lift value of 47.814, as well as associations like "PER KOPLING KAWAHARA 1500RPM" with "PER TORSI KAWAHARA 1500RPM" (lift = 31.893) and "MUR PLONG TITANIUM M8 GR5" with "BAUT TITANIUM GR5 M8X50" (lift = 31.681). These insights not only demonstrate statistically strong relationships but also reflect real-world usage patterns, validated through input from business stakeholders.

The proposed bundling strategies are designed for implementation at the point-of-sale stage, where checkout recommendations can be offered as time-limited discounts or combo deals. This practical approach aligns with actual customer purchasing behavior and the mechanical compatibility of the items, ensuring that both business objectives and customer needs are addressed. With an execution time of just 0.1354 seconds and memory usage of 402.6 MB, the FP-Growth algorithm proved to be both efficient and scalable for retail applications.

The effectiveness of these bundles was further validated through a Customer Satisfaction (CSAT) survey involving 56 respondents, which yielded a high satisfaction score of 83.93%. This score falls into the "Excellent CSAT" category, indicating that customers found the promotional packages relevant, appealing, and valuable. The results affirm that data-driven bundling strategies not only enhance customer satisfaction but also have the potential to increase Average Order Value (AOV) by encouraging additional purchases. Overall, this study supports the use of association rule mining as a practical and profitable tool for small to medium-sized enterprises (SMEs) aiming to boost sales through personalized, evidence-based promotions.

For future research, it is recommended to expand the dataset to cover a full year or more of transaction records to capture seasonal variations and broader purchasing patterns. Additionally, incorporating customer demographic data could allow for more personalized and segment-specific bundling strategies. Comparative analysis with other association rule mining algorithms such as Apriori and Eclat could also provide deeper insights into algorithm performance in terms of runtime, memory usage, and interpretability. Finally, exploring the integration of bundling recommendations into a real-time point-of-sale system and measuring actual AOV impact over time would further validate the practical effectiveness of this approach.

REFERENCES

[1] I. A. A. Alsibhawi, J. B. Yahaya, and H. B. Mohamed, "Business Intelligence Adoption for Small and Medium Enterprises: Conceptual Framework," *Applied Sciences (Switzerland)*, vol. 13, no. 7, Apr. 2023, doi: 10.3390/app13074121.

[2] M. S. Hosen *et al.*, "Data-Driven Decision Making: Advanced Database Systems for Business Intelligence," *Nanotechnol Percept*, vol. 20, no. S3, pp. 687–704, 2024, doi: 10.62441/nano-ntp.v20iS3.51.

[3] M. Yusuf, N. M. Stie, and N. Sangatta, "ANALYSIS OF THE EFFECT OF THE MARKETING MIX ON PURCHASING DECISIONS", [Online]. Available: https://ijemr.politeknikpratama.ac.id/index.php/ijemr

[4] H. Kismo and L. Lahindah, "Designing a Promotion Strategy to Increase Online Sales," *Jurnal Manajemen Bisnis*, vol. 11, no. 2, pp. 1962–1975, Sep. 2024, doi: 10.33096/jmb.v11i2.975.

[5] Q. Wang, X. Zhu, M. Wang, F. Zhou, and S. Cheng, "A theoretical model of factors influencing online consumer purchasing behavior through electronic word of mouth data mining and analysis," *PLoS One*, vol. 18, no. 5 May, May 2023, doi: 10.1371/journal.pone.0286034.

[6] R. M. Teja Nursasongka, I. Fahrurrozi, U. Y. Oktiawati, U. Taufiq, U. Farooq, and G. Alfian, "Utilizing association rule mining for enhancing sales performance in web-based dashboard application," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 2, p. 1105, Nov. 2024, doi: 10.11591/ijeecs.v36.i2.pp1105-1113.

[7] S. Mohammed *et al.*, "A statistical method for predicting quantitative variables in association rule mining," *Inf Syst*, vol. 118, Sep. 2023, doi: 10.1016/j.is.2023.102253.

[8] J. Li, L. Liu, X. Luo, and S. X. Zhu, "Interactive bundle pricing strategy for online pharmacies," *Transp Res E Logist Transp Rev*, vol. 177, Sep. 2023, doi: 10.1016/j.tre.2023.103223.

[9] X. Sun, A. Ngueilbaye, K. Luo, Y. Cai, D. Wu, and J. Z. Huang, "A scalable and flexible basket analysis system for big transaction data in Spark," *Inf Process Manag*, vol. 61, no. 2, Mar. 2024, doi: 10.1016/j.ipm.2023.103577.

[10] A. F. Zakaria, S. C. J. Lim, and M. Aamir, "A pricing optimization modelling for assisted decision making in telecommunication product-service bundling," *International Journal of Information Management Data Insights*, vol. 4, no. 1, Apr. 2024, doi: 10.1016/j.jjimei.2024.100212.

[11] V. Srinadh, "Evaluation of Apriori, FP growth and Eclat association rule mining algorithms," *Int J Health Sci (Qassim)*, pp. 7475–7485, Apr. 2022, doi: 10.53730/ijhs.v6ns2.6729.

[12] R. Wahyuningsih and A. Suharsono, "Comparison Of Market Basket Analysis Method Using Apriori Algorithm, Frequent Pattern Growth (Fp-Growth) And Equivalence Class Transformation (Eclat) (Case Study: Supermarket 'X' Transaction Data For 2021)."

[13] T. A. Djabalul Lael and D. A. Pramudito, "Use of Data Mining for The Analysis of Consumer Purchase Patterns with The Fpgrowth Algorithm on Motor Spare Part Sales Transactions Data," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 4, no. 2, pp. 128–136, Mar. 2023, doi: 10.34306/itsdi.v4i2.582.

[14] A. Wenninger, D. Rau, and M. Röglinger, "Improving customer satisfaction in proactive service design: A Kano model approach," *Electronic Markets*, vol. 32, no. 3, pp. 1399–1418, Sep. 2022, doi: 10.1007/s12525-022-00565-9.

[15] S. K. Jena and A. Ghadge, "Product bundling and advertising strategy for a duopoly supply chain: a power-balance perspective," *Ann Oper Res*, vol. 315, no. 2, pp. 1729–1753, Aug. 2022, doi: 10.1007/s10479-020-03861-9.

[16] N. Purwati, Y. Pedliyansah, H. Kurniawan, S. Karnila, and R. Herwanto, "Komparasi Metode Apriori dan FP-Growth Data Mining Untuk Mengetahui Pola Penjualan," vol. 8, no. 2, 2023.

[17] T. Lakshika and A. Caldera, "Association Rules for Knowledge Discovery From E-News Articles: A Review of Apriori and FP-Growth Algorithms," *Advances in Science, Technology and Engineering Systems Journal*, vol. 7, no. 5, pp. 178–192, 2022, doi: 10.25046/aj070519.

[18] R. Wahyuningsih and A. Suharsono, "Comparison Of Market Basket Analysis Method Using Apriori Algorithm, Frequent Pattern Growth (Fp-Growth) And Equivalence Class Transformation (Eclat) (Case Study: Supermarket 'X' Transaction Data For 2021)."

[19] B. Praveen Kumar, T. Padmavathy, S. U. Muthunagai, and D. Paulraj, "An optimized fuzzy based FP-growth algorithm for mining temporal data," *Journal of Intelligent and Fuzzy Systems*, vol. 46, no. 1, pp. 41–51, Jan. 2024, doi: 10.3233/JIFS-223030.

[20] K. S. Yanisa Putri, I Made Agus Dwi Suarjaya, and Wayan Oger Vihikan, "Sistem Rekomendasi Skincare Menggunakan Metode Content Based Filtering dan Collaborative Filtering," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 3, pp. 764–774, Aug. 2024, doi: 10.51454/decode.v4i3.601.