IJAIT (International Journal of Applied Information Technology)

http://journals.telkomuniversity.ac.id/ijait/

Implementation of the Apriori Algorithm on Outdoor Equipment Rental Transaction Data Based on Clustering Using the K-Means Algorithm

Randi Rizal [a, c], Ruuhwan [b, e], Muhammad Al Husaini [a, *], Dede Rizal Nursamsi [d], Meto Rizki M [b]

[a] *Dept. of Informatic, Siliwangi University, Indonesia*
[b] *Dept. of Informatic, Perjuangan University, Indonesia*
[c] *Faculty of Information and Communication Technology, UTeM, Malaysia*
[d] *Dept. of System Information, Cipasung University, Indonesia*
[e] *Faculty of Informatics, Telkom University, Indonesia*

randirizal@unsil.ac.id, ruuhwan@unper.ac.id, alhusaini@unsil.ac.id, dederizalnursamsi@uncip.ac.id, meto@unper.ac.id

ARTICLE INFO

Received July 25th, 2023
Revised August 23rd, 2023
Accepted August 28th, 2023
Available online Dec. 29th, 2025

Keywords:
Apriori algorithm, K-Means algorithm, data transaction, data mining, association rules, clustering

ABSTRACT

Outdoor equipment rental services play a critical role in meeting climbers' needs prior to expeditions. Sustaining business continuity in this sector requires effective marketing strategies, particularly given the increasing market competition. This study employs data mining techniques to analyze rental transaction data and identify patterns that support strategic decision-making. Specifically, clustering is performed using the K-Means algorithm to group transactions with similar attributes, followed by association rule mining using the Apriori algorithm within each cluster. A dataset comprising 1,276 valid transactions was processed, resulting in three clusters containing 324, 264, and 688 records, respectively, with an accuracy of 0.998. Apriori analysis generated 13 association rules in Cluster 0 and 2 rules in Cluster 1, while no rules met the minimum support and confidence thresholds in Cluster 2 or the overall dataset. These findings demonstrate that clustering prior to association rule mining can uncover meaningful patterns that are not evident in aggregated data. Such insights can inform targeted marketing strategies, including recommendations for item combinations frequently rented together. Future research may integrate alternative algorithms such as ECLAT or FP-Growth and explore framework-based systems to enhance scalability and precision in data-driven decision-making.

Acknowledgment

Thanks to the Department of Informatics, Faculty of Engineering, Siliwangi University, Perjuangan University and Cipasung University for collaborating and providing the opportunity to implement the result of this research.

* Corresponding author at:
Department of Informatics, Faculty of Engineering, Siliwangi University
Jl. Siliwangi No. 24 Kahuripan, Kec. Tawang, Kota Tasikmalaya, 46115, Jawa Barat, Indonesia
E-mail address: alhusaini@unsil.ac.id

ORCID ID:
- First Author: 0000-0002-4377-2137
- Second Author: 0000-0002-0979-0518
- Fourth Author: 0000-0003-2622-0354

http://dx.doi.org/10.25124/ijait.v8i2.6367
Paper_reg_number IJAIT000080203 2024

1. Introduction

Outdoor Equipment Rental is an effort made to meet the needs or equipment to be used by climbers before carrying out climbing activities. Hikers can borrow the equipment needed by renting the equipment within a predetermined time. To ensure the continuity of the rental business, business actors must find solutions and think about accurate marketing strategies. Given the current developments in market conditions, competition between companies in the outdoor equipment rental business is becoming increasingly stringent. One solution to this problem is to analyze transaction data using the concept of data mining so that new conclusions or information can be generated that can be used as a more accurate marketing strategy. The market conditions can be inferred by observing rental data stored in the database, then processed into rental reports and income statements.

In previous research [1][2], there are various methods of data mining to analyze data, including clustering, association rules, classification, and others. The data mining method used to manage transaction data in this research is the association rule method [3]. The association rule technique can provide an overview of rental transaction data patterns that often appear simultaneously in a transaction, so that it will produce a conclusion in the form of a combination of goods that are often rented simultaneously so that later it can be used as marketing information and recommendations. To analyze transaction data that will produce conclusions in the form of an itemset combination, the association rules method using apriori algorithm is utilized [4][5]. The apriori algorithm aims to find frequent itemset that run on a set of data. Thus, apriori analysis is defined as a process to find all apriori rules that meet the minimum requirements for support and minimum requirements for confidence.

However, analyzing market baskets with large itemset tends to be ignored by association rules, thus item recommendations are not precise because there is no information about the product available, resulting in a less precise data obtained. To overcome this, it is necessary to do clustering using the K-Means algorithm to form several groups with the same attributes, then proceed with determining the pattern of association in each group [6][7].

In this research, the authors will implement apriori algorithm on outdoor equipment rental based on clustering using K-Means algorithm. Firstly, existing attributes will be clustered to form the same attribute group, then proceed with determining the association pattern in each clustering group. This will produce association pattern data based on the clustering group and the results of association pattern data without clustering. Thus, it will produce association pattern data to determine more appropriate marketing recommendations.

2. Background and Related Work

2.1. Data Mining

Data Mining, or also known as Knowledge Discovery Database (KDD), is an activity related to data collection, the use of historical data to find knowledge, information, regularity, patterns, or relationships in a large amount of data [8][9]. The output of data mining can be used as an alternative in decision making or to improve the decision making in the future [10][11]. Despite its tremendous potential, data mining faces many challenges related to data privacy, security, and ethical concerns. The abundance of data raises issues about its ownership, consent, and potential misuse. Additionally, biased data or algorithms can lead to unfair outcomes, reinforcing existing inequalities. As data mining continues to evolve, researchers and practitioners are actively working on addressing these challenges while harnessing the power of data to benefit society responsibly.

2.2. Apriori Algorithm

Apriori algorithm is the type of association rules in data mining [12]. It is a process to find all apriori rules that meet the minimum requirements for support and minimum requirements for confidence [13]. Support (support value) is the percentage combination of these items in the database, while confidence (certainty value) is the strength of the relationship between items in the association rules.

2.2.1. High Frequency Pattern Analysis with Apriori Algorithm

This stage looks for item combinations that meet the requirements of the support value in the database. The support value of an item is obtained by using Equation 1.

$$Support\ (A) = \frac{\sum Transaction\ contain\ A}{\sum Transaction} \qquad \text{Equation 1}$$

Meanwhile, the support value of the 2 items is obtained by using Equation 2 and 3.

$$Support\ (A, B) = P\ (A \cap B) \qquad \text{Equation 2}$$

$$Support\ (A) = \frac{\sum Transaction\ contain\ A,B}{\sum Transaction} \qquad \text{Equation 3}$$

Frequent item shows an itemset that have a frequency of occurrence that is more than the specified minimum value of (∅). For example, ∅=3 then all itemset whose frequency of occurrence is ≥ 3.

2.2.2. Formation of Association Rules

After all high-frequency patterns are found, then measures association rules that meet the minimum requirements for the confidence value by calculating the confidence of the A_B rules with Equation 4 and 5.

$$Confidence\ P = (B|A) \qquad \text{Equation 4}$$

$$Confidence = \frac{\sum Transaction\ contain\ A,B}{\sum Transaction} \qquad \text{Equation 5}$$

To determine the association rules to be selected, it must be derived based on support x confidence.

2.3. K-Means Algorithm

K-Means algorithm is a non-hierarchical data clustering method that seeks to partition existing data into one or more clusters so that data with the same characteristics are grouped into the same cluster and data with different characteristics are grouped into other groups [14]. K-Means algorithm basically performs two processes, namely the process of detecting the location of the center of each cluster and the process of searching for members from each cluster [15]. The stages of K-Means algorithm are as follows:

1. Specify K as the number of formed clusters.
2. Generate the initial K centroid (cluster center point) randomly.
3. Calculate the distance of each data to each centroid with the Euclidean formula in Equation 6.

$$D(a,b) = \sqrt{\sum_{k=1}^{n}(ak - bk)^2} \quad \text{Equation 6}$$

4. Calculate data to the nearest centroid.
5. Determine the new centroid position by calculating the average value of the data located at the same centroid.
6. Return to step three (3) if the position of the new centroid is not the same as the old centroid.

2.4. Related Work

Based on research [16], Apriori algorithm on accessory product sales data was implemented. It produced six rules that can be used as material for a sales strategy by pairing the first product with the second product by providing an attractive price. Study [17] implements K-Means algorithm to classify the vulnerability of Rabies cases using the CRISP-DM methodology through the process of understanding business, understanding data, preparing data, modeling, evaluating, and disseminating the data. In addition, in [18], K-Means algorithm has also been used for grouping patient data exposed to the Covid 19 virus. Moreover, implementation of apriori algorithm in research [4] and [19] is used to find information on patterns of the most sales data and the relationship between one product and another [20][21][22].

3. Methodology

The flow of the research methodology is illustrated in Figure 1.

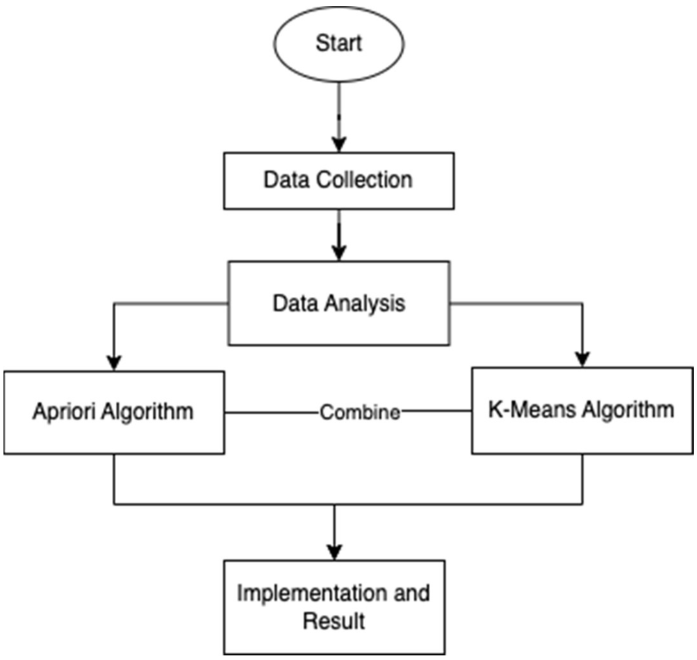


Figure 1 Research Methodology

3.1. Data Collection

Astrajingga Adventure is the most popular outdoor equipment rental place for customers in Tasikmalaya. Based on the results of interviews conducted by the author, Astrajingga Outdoor provides various rental equipment which are dome with a capacity of 2-5 people, scout ropes, fly sheets and backpacks, sleeping bags, mattresses, headlamps, tent lights, shoes, stoves, nesting, and portable gas. There are 1298 transaction data collected in a year, and data selection was carried out to classify 1276 valid data and 22 invalid data. Invalid data is transaction data that fails or is not executed between tenants and customers. Table 1 summarizes transaction data that has been collected.

Table 1 Data Transaction

No.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1.	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
2.	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
3.	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4.	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0
5.	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
6.	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
7.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
8.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9.	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0
10.	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
….	….	….	….	….	….	….	….	….	….	….	….	….	….	….	….
1276	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0

Note: A (Dome 5), B (Dome 4), C (Dome 2), D (Scout ropes), E (Fly sheet 3x3), F (Fly sheet 3x6), G (Backpacks), H (Sleeping Bag), I (Mattresses), J (Headlamps), K (Tent light), L (Shoes), M (Stoves), N (Nesting), O (Portable gas).

4. Results and Discussion

4.1. Data Analysis

At this stage, the author analyzes the transaction data that has been collected using two methods or algorithms to get the appropriate results.

4.1.1. Grouping Transaction Data with The K-Means Algorithm

At this stage, transaction data is grouped using K-Means algorithm with the help of the RapidMiner application as a tool to analyze 1276 transaction data. Figure 2 illustrates the cluster data count performance operator to assess the accuracy of the number of clusters to be created.

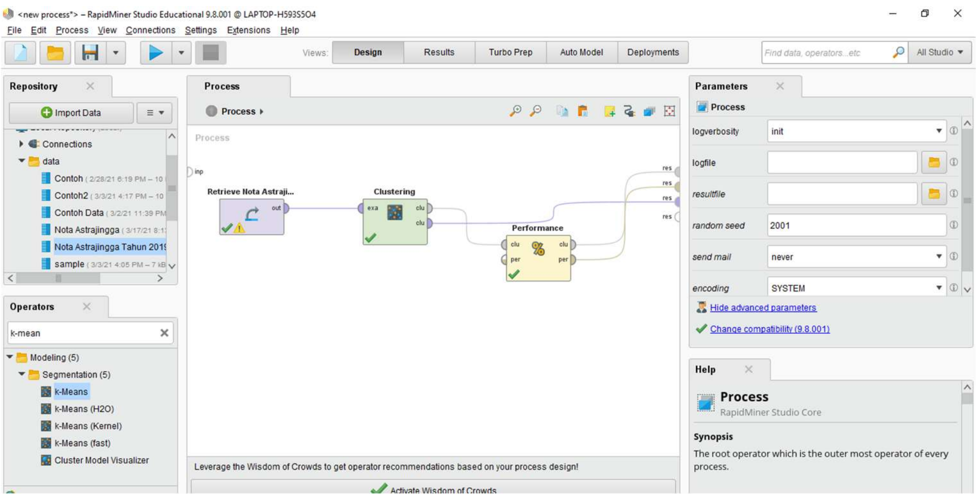
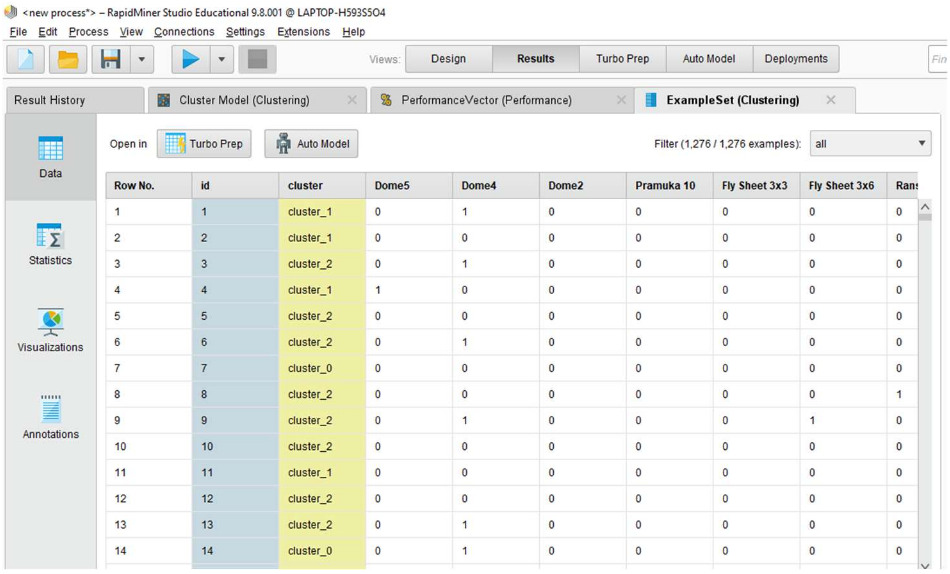


Figure 2 Data Display and Connected Operators

After starting the application, the results display will appear, namely there are clustering results in the form of tables as shown in Figure 3, performance results shown in Figure 4, and the number of items in each cluster shown in Figure 5.



Row No.	id	cluster	Dome5	Dome4	Dome2	Pramuka 10	Fly Sheet 3x3	Fly Sheet 3x6	Rans
1	1	cluster_1	0	1	0	0	0	0	0
2	2	cluster_1	0	0	0	0	0	0	0
3	3	cluster_2	0	1	0	0	0	0	0
4	4	cluster_1	1	0	0	0	0	0	0
5	5	cluster_2	0	0	0	0	0	0	0
6	6	cluster_2	0	1	0	0	0	0	0
7	7	cluster_0	0	0	0	0	0	0	0
8	8	cluster_2	0	0	0	0	0	0	1
9	9	cluster_2	0	1	0	0	0	1	0
10	10	cluster_2	0	0	0	0	0	0	0
11	11	cluster_1	0	0	0	0	0	0	0
12	12	cluster_2	0	0	0	0	0	0	0
13	13	cluster_2	0	1	0	0	0	0	0
14	14	cluster_0	0	1	0	0	0	0	0

Figure 3 Display Results in The Form of a Table

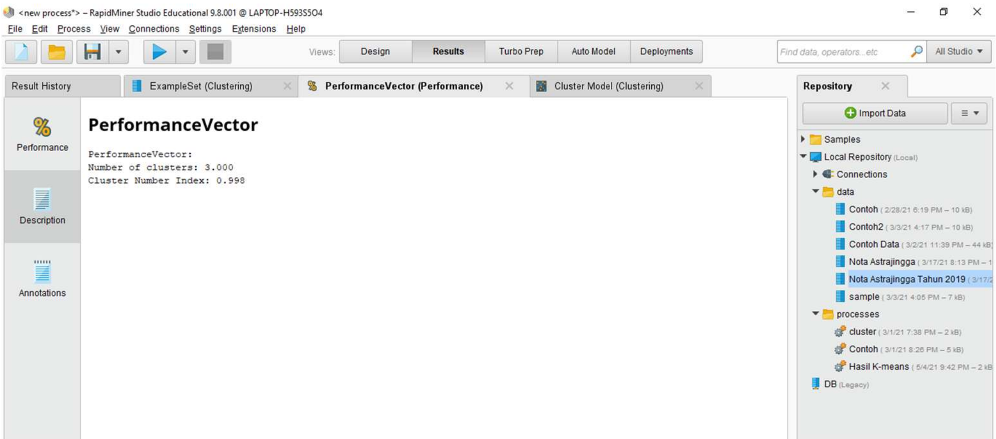


Figure 4 Display Performance

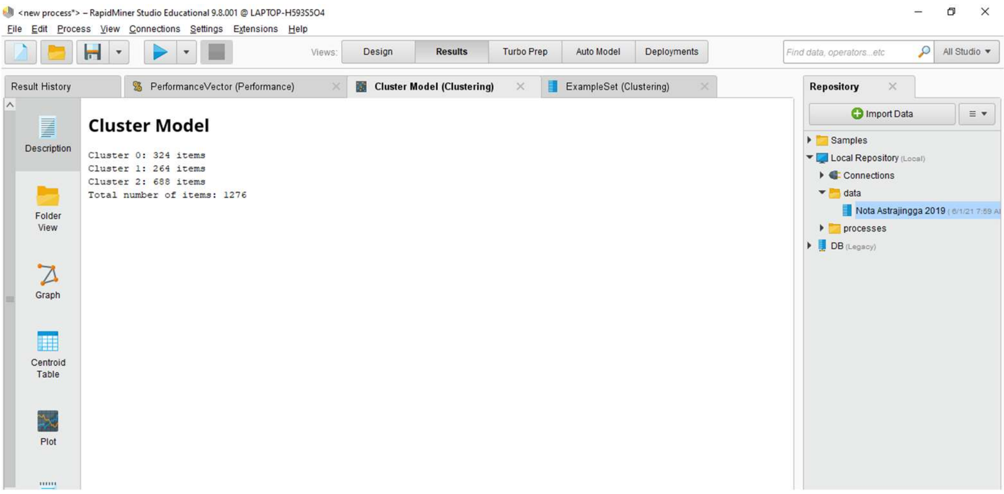


Figure 5 Display the Number of Items in Each Cluster.

Looking at the results of data calculations using K-Means algorithm with the number of clusters being 3, the results obtained are cluster 0 = 324 items, cluster 1 = 264 items, and cluster 2 = 688 items with an accuracy level of 0.998.

Implementation of the calculation of K-Means manually with the cluster to be formed is k = 3 based on the previous accuracy value which suggests that dividing the data into three clusters would be an appropriate approach. This manual process involves iteratively adjusting cluster centroids until an optimal solution is reached, aligning the clusters with the data's inherent patterns as closely as possible, as shown in Table 2.

Table 2 Determine the Initial Centroid

No.	Item	Cluster 0	Cluster 1	Cluster 2
1.	Dome Capacity 5	0.272	0.292	0.257
2.	Dome Capacity 4	0.392	0.364	0.323
3.	Dome Capacity 2	0.080	0.072	0.097
4.	Scout ropes	0	0	0.003
5.	Fly Sheet 3x3	0.173	0.133	0.093
6.	Fly Sheet 3x6	0.040	0.121	0.103
7.	Backpackers 60-80 L	0.191	0.295	0.170
8.	Sleeping Bag	0.228	0.345	0.192
9.	Mattresses	0.451	1	0
10.	Headlamp	0.256	0.159	0.119
11.	Tent light	0.003	0	0.001
12.	Shoes	0.080	0.083	0.172
13.	Stoves	0.873	0.034	0.071
14.	Nesting	0.719	0.064	0.031
15.	Portable gas	0.877	0.110	0.115

Calculating the distance to the first transaction data with the center point (centroid) in cluster 0.

1. Calculation on the first data

$$
\begin{aligned}
D(1,0) = \sqrt{(0-0.272)^2} &+ (1-0.392)^2 + (0-0.080)^2 + (0-0)^2 \\
&+ (0-0.173)^2 + (0-0.040)^2 + (0-0.191)^2 \\
&+ (0-0.228)^2 + (1-0.451)^2 + (0-0.256)^2 \\
&+ (0-0.003)^2 + (0-0.080)^2 + (0-0.873)^2 \\
&+ (0-0.719)^2 + (0-0.877)^2 = 1.729
\end{aligned}
$$

2. Calculation on the cluster 1

$$
\begin{aligned}
D(1,1) = \sqrt{(0-0.292)^2} &+ (1-0.364)^2 + (0-0.072)^2 + (0-0)^2 \\
&+ (0-0.133)^2 + (0-0.121)^2 + (0-0.295)^2 \\
&+ (0-0.345)^2 + (1-1)^2 + (0-0.159)^2 \\
&+ (0-0)^2 + (0-0.083)^2 + (0-0.034)^2 \\
&+ (0-0.064)^2 + (0-0.110)^2 = 0.884
\end{aligned}
$$

3. Calculation on the cluster 2

$$
\begin{aligned}
D(1,2) = \sqrt{(0-0.257)^2} &+ (1-0.323)^2 + (0-0.097)^2 \\
&+ (0-0.003)^2 + (0-0.093)^2 + (0-0.103)^2 \\
&+ (0-0.170)^2 + (0-0.192)^2 + (1-0)^2 \\
&+ (0-0.119)^2 + (0-0.001)^2 + (0-0.172)^2 \\
&+ (0-0.071)^2 + (0-0.031)^2 + (0-0.115)^2 \\
&= 1.296
\end{aligned}
$$

Table 3 is the result of calculating transaction data with a center point in iteration 1 using Euclidean distances.

Table 3 The Result of Calculating the Distance of The Data to The Center Point in Iteration 1

No.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	C0	C1	C2	Cluster
1.	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1.729	0.884	1.296	1
2.	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1.666	0.714	1.152	1
3.	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1.701	1.335	0.825	2
4.	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1.943	1.112	1.558	1
5.	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1.779	1.480	1.043	2

No.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	C0	C1	C2	Cluster
6.	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1.701	1.335	0.825	2
7.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1.085	1.795	1.398	0
8.	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1.815	1.385	0.993	2
9.	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	1.751	1.863	1.527	2
….	….	….	….	….	….	….	….	….	….	….	….	….	….	….	….	…	…	…	…
1276	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	2.212	1.881	1.553	2

From Table 3, it can be inferred that the number of items in each temporary cluster is cluster 0 = 324, cluster 1 = 264, and cluster 2 = 688.

4.1.2. Search for Transaction Data Association Patterns in Each Cluster with the Apriori Algorithm

At this stage, a search for patterns of transaction data associations will be carried out in each cluster using the Apriori algorithm with the provision that the minimum support value is ≥ 30% common itemset are prioritized and the minimum confidence value is ≥ 75%, only strong associations are considered. RapidMines is used to find association patterns on transaction data in each cluster. The process operators discretize by frequency, numerical to binomial, fp-growth, and create association rules as shown Figure 6.

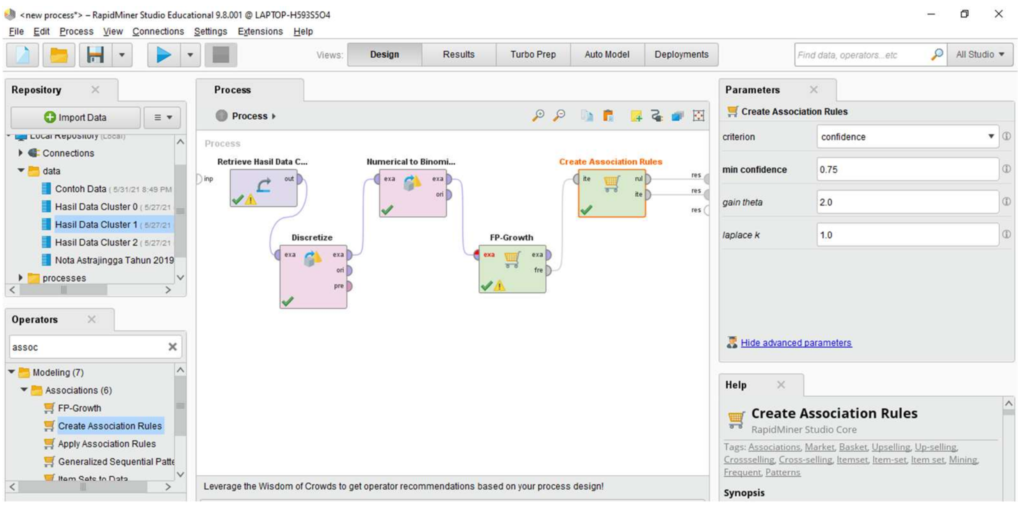


Figure 6 Processing Data Using Apriori

The results of the rules that are formed in each data cluster with the apriori algorithm are as follows.

1. Search for Association Patterns in Cluster 0. Formation of a combination of 1 itemset on cluster 0 transaction data with a minimum support value of 30%. Table 4, Table 5, Table 6, Table 7 are the results of the support value for 1 itemset combination that meets the minimum support value, as calculated using the formula in Equation 7.

$$Support\ (A) = \frac{\sum Transaction\ contain\ A}{\sum Transaction} \qquad \text{Equation 7}$$

$$Support\ (Gas) = \frac{284}{324} = 0.8765$$

Table 4 The Value of Support in The Combination of 1 Itemset in Cluster 0

No.	Name Item	$\sum$ Transaction	Support	
1.	Gas Portable	284	87.65%	0.876543
2.	Stoves	283	87.35%	0.873457
3.	Nesting	233	71.91%	0.719136
4.	Mattresses	146	45.06%	0.450617
5.	Dome 4	127	39.20%	0.391975

Then, combine two itemset in Table 4 and then determine the support value using the formula provided in Equation 8 for each of the 2 itemset.

$$Support\ (A, B) = P\ (A \cap B) \qquad \text{Equation 8}$$

$$Support\ (Stoves, Gas\ Portable) = \frac{243}{324} = 0.75$$

Table 5 The Value of Support in The Combination of 2 Itemset in Cluster 0

No.	Name Item	$\sum$ Transaction	Support	
1.	Stoves, Gas Portable	243	75.00%	0.75
2.	Gas Portable, Nesting	193	59.57%	0.595679
3.	Stoves, Nesting	192	59.26%	0.592593
4.	Stoves, Matras	126	38.89%	0.388889
5.	Nesting, Matras	125	38.58%	0.385802
6.	Gas Portable, Matras	124	38.27%	0.382716
7.	Stoves, Dome 4	116	35.80%	0.358025
8.	Gas Portable, Dome 4	110	33.95%	0.339506
9.	Nesting, Dome 4	90	27.78%	0.277778
10.	Mattresses, Dome 4	67	20.68%	0.206790

Afterward, combine 3 itemset in the previous table and then determine the support value using the formula provided in Equation 9 for each of the 3 itemset.

$$Support\ (A, B, C) = P\ (A \cap B \cap C) \qquad \text{Equation 9}$$

$$Support\ (Stoves, Gas\ Portable, Nesting) = \frac{152}{324} = 0.469136$$

Table 6 The Value of Support in The Combination of 3 Itemset in Cluster 0

No.	Name Item	$\sum$ Transaction	Support	
1.	Stoves, Gas Portable, Nesting	152	46.91%	0.469136
2.	Stoves, Nesting, Matras	105	32.41%	0.324074
3.	Stoves, Gas Portable, Matras	104	32.10%	0.320988
4.	Gas Portable, Nesting, Matras	103	31.79%	0.317901
5.	Stoves, Gas Portable, Dome 4	99	30.56%	0.305556
6.	Stoves, Nesting, Dome 4	79	24.38%	0.243827
7.	Gas Portable, Nesting, Dome 4	73	22.53%	0.225309
8.	Stoves, Matras, Dome 4	61	18.83%	0.188272
9.	Nesting, Matras, Dome 4	56	17.28%	0.17284
10.	Mattresses, Dome 4	67	20.68%	0.206790

In cluster 0 data experiment, there are 13 rules formed as presented in Table 7.

Table 7 The Results of The Rules on Data Cluster 0

No.	If (A)	Then (B)	$\sum$ Transaction (A, B)	$\sum$ Transaction (A)	Support (%)	Confidence	Lift Ratio
1.	Stoves	Gas Portable	243	283	75.00	85.87%	1.144
2.	Nesting	Gas Portable	193	233	59.57	82.83%	1.390
3.	Nesting	Stoves	192	233	59.26	82.40%	1.390
4.	Stoves, Nesting	Gas Portable	152	192	46.91	79.17%	1.687
5.	Matras	Stoves	126	146	38.89	86.30%	2.219
6.	Matras	Nesting	125	146	38.58	85.62%	2.219
7.	Matras	Gas Portable	124	146	38.27	84.93%	2.219

No.	If (A)	Then (B)	∑ Transaction (A, B)	∑ Transaction (A)	Support (%)	Confidence	Lift Ratio
8.	Dome 4	Stoves	116	127	35.80	91.34%	2.551
9.	Dome 4	Gas Portable	110	127	33.95	86.61%	2.551
10.	Nesting, Matras	Stoves	105	125	32.41	84.00%	2.592
11.	Gas Portable, Matras	Stoves	104	124	32.10	83.87%	2.612
12.	Gas Portable, Matras	Nesting	103	124	31.79	83.06%	2.612
13.	Gas Portable, Dome 4	Stoves	99	110	30.56	90.00%	2.945

2. Search for Association Patterns in Cluster 1. Formation of a combination of 1 itemset on cluster 1 transaction data with a minimum support value of 30%. Table 8, Table 9, Table 10, Table 11 are the results of the support value for the 1 itemset combination that meets the minimum support value.

$$Support\ (Mattresses) = \frac{264}{264} = 1$$

Table 8 The Value of Support in The Combination of 1 Itemset in Cluster 1

No.	Name Item	$\sum$ Transaction	Support	
1.	Mattresses	264	100.00%	1
2.	Dome 4	96	36.36%	0.363636
3.	Sleeping Bag	91	34.47%	0.344697

Next, combine 2 itemset in Table 9 and then determine the support value for each of the 2 itemset.

$$Support\ (Mattresses, Dom\ 4) = \frac{96}{264} = 0.363636$$

Table 9 The Value of Support in The Combination of 2 Itemset in Cluster 1

No.	Name Item	$\sum$ Transaction	Support	
1.	Mattresses, Dome 4	96	36.36%	0.363636
2.	Mattresses, Sleeping Bag	91	34.47%	0.244697
3.	Dome 4, Sleeping Bag	25	9.47%	0.094697

Next, combine 3 itemset in Table 9 and then determine the support value for each of the 3 itemset.

$$Support\ (Mattresses, Dome\ 4, Sleeping\ Bag) = \frac{25}{264} = 0.09469$$

Table 10 The Value of Support in The Combination of 3 Itemset in Cluster 1

No.	Name Item	$\sum$ Transaction	Support	
1.	Mattresses, Dome 4, Sleeping Bag	25	9.47%	0.094697

In the cluster 0 data experiment, there are 2 rules formed as presented in Table 11.

Table 11 The Results of The Rules on Data Cluster 1

No.	If (A)	Then (B)	∑ Transaction (A, B)	∑ Transaction (A)	Support (%)	Confidence (%)	Lift Ratio
1.	Dome 4	Mattresses	96	96	36.36	100.00	2.75
2.	Sleeping Bag	Mattresses	91	91	34.47	100.00	2.90

3. Search for Association Patterns in Cluster 2. In the Cluster 2 data experiment, there were no rules formed that met the specified minimum support and minimum confidence values.

4.1.3. Comparison of Rules Results

Figure 7 is a comparison graph of the results of the rules formed on the data from each cluster with the entire data.

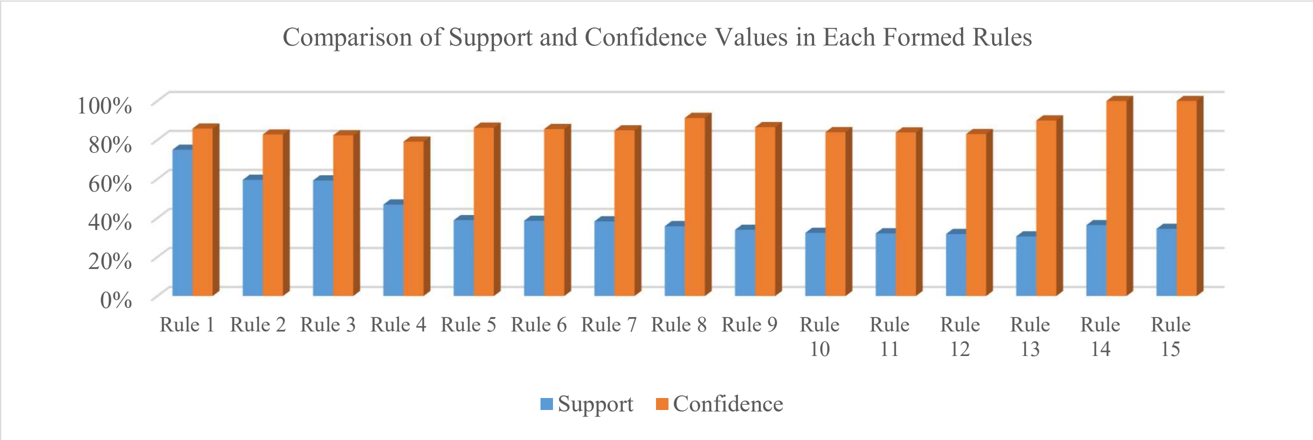


Figure 7 Comparison of Support and Confidence Values in Each Formed Rule

From the results, data in cluster 0 and 1 yields rule, while the data in cluster 2 and the entire data do not find rules that meet predetermined values. The outcome of a rule generation or analysis process conducted on clustered data. Rules meeting predetermined criteria were found for data in Cluster 0 and Cluster 1, but no such rules were identified for data in Cluster 2 or the entire dataset. This result has implications for understanding the structure and characteristics of the data within different clusters.

5. Conclusions

Implementation of the K-Means algorithm for grouping outdoor equipment rental transaction data with a total of 1276 transaction data obtained results with a total of 3 clusters with an accuracy level of 0.998 with details of the amount of data in cluster 0 = 324, cluster 1 = 264, and cluster 2 = 688 data. For the implementation of the apriori algorithm, it produces 13 rules in cluster 0 and 2 rules in cluster 1. However, there are no rules found in cluster 2 and overall data that meet the minimum support and confidence requirements. For future research in grouping and searching for association patterns, it can be combined with several other algorithms in association rules, such as ECLAT and FP-Growth or added with other methods in data mining. In addition, grouping analysis and association pattern search can be implemented in the form of a framework-based system so that it can analyze data.

Bibliography

[1] J. Sun and Y. Sun, "Analysis of Data Mining Based on Object Oriented Analysis Method," *Journal of Physics: Conference Series*, vol. 1069, p. 012041, Aug. 2018. DOI: 10.1088/1742-6596/1069/1/012041

[2] X. Shu and Y. Ye, "Knowledge Discovery: Methods from data mining and machine learning," *Social Science Research*, vol. 110, p. 102817, Feb. 2023. DOI: 10.1016/j.ssresearch.2022.102817

[3] A. Lia Hananto *et al.*, "Analysis of Drug Data Mining With Clustering Technique Using K-Means Algorithm," *Journal of Physics: Conference Series*, vol. 1908, no. 1, p. 012024, Jun. 2021. DOI: 10.1088/1742-6596/1908/1/012024

[4] R. Alfianzah, R. I. Handayani, and M. Murniyati, "Implementation of Apriori Algorithm Data Mining for Increase Sales," *SinkrOn*, vol. 5, no. 1, p. 17, Oct. 2020. DOI: 10.33395/sinkron.v5i1.10587

[5] A. H. Nasyuha, Z. Zulham, and I. Rusydi, "Implementation of K-means algorithm in data analysis," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 2, p. 307, Apr. 2022. DOI: 10.12928/telkomnika.v20i2.21986

[6] E. T. Munawaroh, R. I. Handayani, and E. Widanengsih, "The Implementation of K-Means Algorithm for Cluster Majoring to New Students in SMKN 2 of South Tangerang," *SinkrOn*, vol. 4, no. 1, p. 164, Oct. 2019. DOI: 10.33395/sinkron.v4i1.10133

[7] Y. Mulyana, A. Fadlil, and I. Riadi, "Classification of Organizational Data Using the K-Means Algorithm," *Komputasi: Jurnal Ilmiah Ilmu Komputer dan Matematika*, vol. 19, no. 2, pp. 123–137, Jul. 2022. DOI: 10.33751/komputasi.v19i2.5676

[8] A. M. Abdulkadium, R. A. A. Shekan, and H. A. Hussain, "Application of Data Mining and Knowledge Discovery in Medical Databases," *Webology*, vol. 19, no. 1, pp. 4912–4924, Jan. 2022. DOI: 10.14704/WEB/V19I1/WEB19329

[9] F. A. Rahman, M. I. Desa, A. Wibowo, and N. A. Haris, "Knowledge Discovery Database (KDD)-Data Mining Application in Transportation," *Proceeding of the Electrical Engineering Computer Science and Informatics*, vol. 1, no. 1, Aug. 2014. DOI: 10.11591/eecsi.v1.357

[10] H. Mydyti, "Data Mining Approach Improving Decision-Making Competency along the Business Digital Transformation Journey: A Case Study – Home Appliances after Sales Service," *SEEU Review*, vol. 16, no. 1, pp. 45–65, Jul. 2021. DOI: 10.2478/seeur-2021-0008

[11] S. Dash, S. K. Shakyawar, M. Sharma, and S. Kaushik, "Big data in healthcare: management, analysis and future prospects," *Journal of Big Data*, vol. 6, no. 1, p. 54, Dec. 2019. DOI: 10.1186/s40537-019-0217-0

[12] H. Wu, "Data Association Rules Mining Method Based on Improved Apriori Algorithm," in *2020 the 4th International Conference on Big Data Research (ICBDR'20)*, 2020, pp. 12–17. DOI: 10.1145/3445945.3445948

[13] E. H. Hermaliani, L. Kurniawati, T. Haryanti, N. Mutiah, A. Kurniawan, and B. S. Renhoran, "Data Mining Technique to Determine the Pattern of Fruits Sales & Supplies Using Apriori Algorithm," *Journal of Physics: Conference Series*, vol. 1641, no. 1, p. 012070, Nov. 2020. DOI: 10.1088/1742-6596/1641/1/012070

[14] N. A. Khairani and E. Sutoyo, "Application of K-Means Clustering Algorithm for Determination of Fire-Prone Areas Utilizing Hotspots in West Kalimantan Province," *International Journal of Advances in Data and Information Systems*, vol. 1, no. 1, pp. 9–16, Apr. 2020. DOI: 10.25008/ijadis.v1i1.13

[15] C. Prianto, R. Nuraini, and A. T. Wali, "Implementation of K-Means Methods In Clustering Students Ability Levels in English Language," *The IJICS (International Journal of Informatics and Computer Science)*, vol. 3, no. 2, p. 49, Oct. 2019. DOI: 10.30865/ijics.v3i2.1382

[16] S. Qomariah, B. Basrie, and S. F. Pa'a, "Implementasi Algoritma Apriori Pada Data Penjuanlan Produk Asesoris Cv Princes Diary Samarinda," *Just TI (Jurnal Sains Terapan Teknologi Informasi)*, vol. 12, no. 2, p. 31, Jul. 2020. DOI: 10.46964/justti.v12i2.321

[17] K. Rahayu, L. Novianti, and M. Kusnandar, "Implementation Data Mining With K-Means Algorithm For Clustering Distribution Rabies Case Area In Palembang City," *Journal of Physics: Conference Series*, vol. 1500, no. 1, p. 012121, Apr. 2020. DOI: 10.1088/1742-6596/1500/1/012121

[18] P. Mas, A. Putra, I. Gusti, A. Gede, and A. Kadyanan, "Implementation of K-Means Clustering Algorithm in Determining Classification of the Spread of the COVID-19 Virus in Bali," *Jurnal Elektronik Ilmu Komputer Udayana*, vol. 10, no. 1, 2021.

[19] R. Fenny Syafariani, “Association Analysis with Apriori Algorithm for Electronic Sales Decision Support System,” *International Journal of Informatics Information System and Computer Engineering*, vol. 3, no. 1, pp. 61–72, 2022 [Online]. DOI: 10.34010/injiiscom.v3i1.8089

[20] A. W. O. Gama and N. M. Widnyani, “Simple Modification for an Apriori Algorithm With Combination Reduction and Iteration Limitation Technique,” *Knowledge Engineering and Data Science*, vol. 3, no. 2, p. 89, Dec. 2020. DOI: 10.17977/um018v3i22020p89-98

[21] A. I. Idris *et al.*, “Comparison of Apriori, Apriori-TID and FP-Growth Algorithms in Market Basket Analysis at Grocery Stores,” *The IJICS (International Journal of Informatics and Computer Science)*, vol. 6, no. 2, p. 107, Jul. 2022. DOI: 10.30865/ijics.v6i2.4535

[22] J. Mantik, A. Almayda, and S. Saepudin, “Application of K-Means Clustering Data Mining in the Extension of the Contract of Work at PT. Afixkogyo Indonesiaid 2,” 2021.