Journal of Computer Science and Information Technology
https://jcsitech-upiyptk.org/ojs
20 2 2 Volumes 8 Issue 4 Matter: 98-105 e-ISSN: 2502-1486

Customer Grouping for Customer Relationship Management Optimization with the K-Means Algorithm

Zidane Dwi Montero [1]
[1] Putra Indonesia University YPTK Padang
zidanedwimontero@gmail.com

Abstract

Customer Relationship Management (CRM) is a tool that can help organizations/companies to achieve their goals. Models that can be developed in relationships are trust and commitment. Trust can come when both parties share experiences. The rapid development of information and technology has led to significant changes in business competition. Rani 2 Supermarkets are modern shopping centers that provide a variety of household needs at low prices and quality. Rani Supermarkets is located at Jl. Zeinizein Painan Pesisir Selatan District. The Rani 2 Supermarket has implemented strategies to build customer loyalty, such as providing additional discounts, facilitating transactions, giving certain customer cards. The problem is that supermarkets don't know which customers are loyal and haven't based it on a pattern. One method that can be used in analyzing CRM is data mining. Clustering with the K-Means grouping technique will produce customer information based on clusters. This cluster is divided into two, namely productive customer clusters and less productive customer clusters. The problem to be solved with this method is how knowledge is useful for the company, namely increasing customer loyalty. Through the K-Means Cluster Technique, this will be used as material for implementing a customer relationship management strategy with the aim of increasing customer loyalty and sales volume.

Keywords : Customer Relationship Management, Loyalty, K-Means Algorithm, Clustering, Method .

1. Introduction

Customer Relationship Management (CRM) is a tool that can help organizations/companies to achieve their goals. The goal is to effectively and efficiently improve customer acquisition and retention. Customer Relationship Management (CRM) is used to describe a variety of software applications used to optimize marketing, sales and customer service functions. Service quality is the basis for successful CRM strategy implementation to produce the performance expected by customers. This will improve the relationship with customers which is a functional strategy to create share value. [1].

In the current era of globalization, the rapid development of information and technology has led to significant changes in business competition in every company. In this case, technology and information are very important at this time because they can help work or activities.

Information System is a system that is very important for companies in an organization that meet the needs of daily transaction processing. As is the case with Rani Supermarkets, which have implemented strategies to build customer loyalty, such as providing additional discounts, facilitating transactions, giving certain customer cards.

Supermarket Rani 2 is a modern shopping center that provides a variety of household needs that come with low prices and quality. Rani Supermarket is located at Jl. Zeinizein Painan Pesisir Selatan District. The Rani 2 Supermarket has implemented strategies to build customer loyalty, such as providing additional discounts, facilitating transactions, giving certain customer cards.

The problem at Rani Supermarket 2 is that supermarkets don't know which customers are loyal to the products provided and the company hasn't based it on a specific customer pattern. This problem can be overcome by building a Customer Relationship Management (CRM). With this, it can foster customer loyalty , foster long-term relationships to create greater value so as to be able to maintain market share and customer loyalty [2].

Much research has been done on customer grouping with the concept of Customer Relationship Management and the K-Means Algorithm Method [3] conducting research using K-Means Clustering to determine the best and potential customers from the Medan Post Office where it is found that the best customers are obtained from customers with the largest number of customers. a lot of transactions and a moderate or high amount of money.

Furthermore, the second [4] conducts research also applies the K-Means Clustering algorithm to find out

Received: 10-09-2022 | Revision: 22-09-2022 | Published: 31-10-2022 | doi: 10.35134/jcsitech.v8i4.46

potential customers at PT Sinar Kencana Intermoda Surabaya. In this research, data analysis is carried out in two ways, namely RFM weighting to produce RFM weights where recency is the final transaction, frequency is the number of transactions and monetary is the number of transactions. After that, they were grouped using the K-Means method. From the results of the system evaluation it was found that the grouping of customers in the BZ category had a percentage value of 54.3 % , the MVC category was 21.8% and the MGC was 23.9.

One method that can be used in analyzing CRM is data mining. Clustering with the K-Means grouping technique will produce customer information based on clusters. This cluster is divided into two, namely productive customer clusters and less productive customer clusters. The problem to be solved with this method is how knowledge is useful for the company, namely increasing customer loyalty. Through the K-Means Cluster Technique, this will be used as material for implementing a customer relationship management strategy with the aim of increasing customer loyalty and sales volume.

According to Philip Kotler and Kevin Lane Keller [5] quoted from the book Marketing Management, customer satisfaction is a person's feeling of pleasure or disappointment that arises after comparing the expected service performance (outcome) to the expected performance. Satisfying consumer needs is a desire in every company.

K-Means is one of the well-known algorithms in clustering, originally known as Forgy's and has been used extensively in various fields including data mining, statistical data analysis and other business applications [6]. K-Means Clustering is a local optimization method that is sensitive to selecting the initial position of the cluster midpoint. So choosing a bad initial position from the midpoint of the cluster will result in the K-Means Clustering algorithm being stuck in the optimal local solution [7].

According to [8] the web is divided into two, namely static web and dynamic web, static web is web content that cannot change, meaning that the contents of documents cannot be changed, or cannot update the contents. The technology used for static web is a type of CSS such as HTML, an example of a static website is a company's web profile that predominantly uses HTML.

PHP is server-side programming, which is a programming language that is processed on the server side. According to Supono [9] " PHP (Hypertext Preprocessor) is a programming language used to translate lines of program code into server-side understandable machine code that can be added to HTML.

SQL acts as a language that regulates data transactions between applications and databases as data storage. With SQL, network and application programmers have no difficulty at all in connecting the applications they make [10]. Commonly used databases include MYSQL, Oracle, SQL server and so on

2. Research methodology

To assist in the preparation of this research so that the steps in solving the problems to be discussed can be clearly structured, it is necessary to have a framework arrangement. The research framework contained in Figure.1.

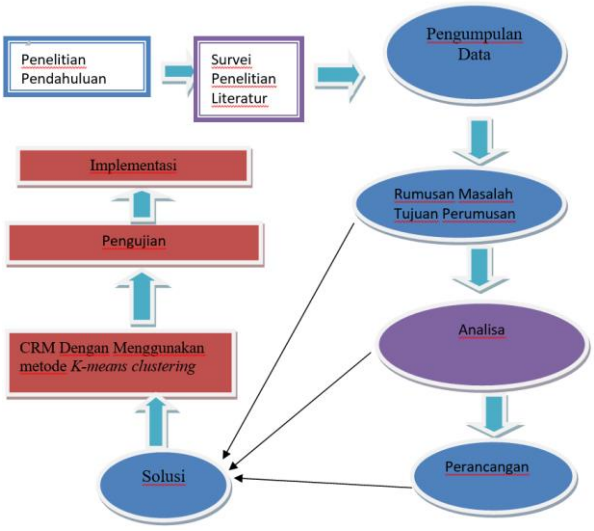

Figure 1. Research Framework

The research stage is a sequence of processes or steps that will be carried out in completing this research. The stages of this research are as follows:

2.1 Field Research

This research was conducted with the owner or manager of Rani Swalayan 2. To obtain data about Rani Swalayan 2, and request data and other necessary information .

2.2 Research Library (Library Research)

Data collection on Rani Swalayan 2 in Pesisir Selatan District was carried out by reading reference books, journals, and searching the internet. So that the data obtained can be used as a basis for the next research stage.

2.3 Analysis

Based on the problem identification above, the researcher conducted data analysis first. This is so problem solving can produce new solutions.

2.4 System planning

At the design stage of the Customer Relationship Management system with the K-Means Algorithm

using the PHP programming language and MySQL database.

2.5 System Implementation

The testing method used in this study is the direct testing method, namely by using interface testing. Used to test the special functions of the designed software.

3. Results and Discussion

3.1 K-Means Algorithm

In the early stages the data will be processed using the a priori algorithm method. The following is a product list table.

Table 1. Product Data Table

id	Name of goods	Number of Transactions	Total Transactions
001	Solar Salt Warehouse 16	4	102,000
002	Magnum mild 16 smp	1	21,000
003	Tea bottle	1	6,000
004	Shinzui kirei soap	1	4,000
005	Detergent	1	5,000
006	Biscuits	1	8,000
007	Ademsari	3	6,000
008	Sampoerna 16 btg mild	1	25,500
009	Fried Indomie	5	13,000
010	Choki-choki	3	3,000
011	Merries 11 S	2	38,000
012	Blue band 200 ml gr	2	22,000
013	SoyaMaster 300 ml	1	10,000
014	Walls paddle pop choco	1	2,000
015	Chitato	1	6,000
016	Magnum Mild 16 junior high school	3	63,000
017	Cimory 120 ml of strawberries	1	11,000
018	Pandan Wangi Rice	1	135,000
019	Bobo Slay	3	9,000
020	Six Star	1	10,500
021	Marjan Lemon	1	22,000
022	SoyaMaster 200 ml	2	10,000
023	Bobo Sweet Bread	3	24,000
024	Roma Sari Wheat 249 gr	1	9,000
025	Aqua Gallon	1	20,000
026	Golda coffee	3	9,000
027	cricket	1	3,500
028	Rock sugar Rio tea	5	5,000
029	Talua Beans	1	11,000
030	Honey TJ	8	16,000

In this case the author took data at Rani Swalayan 2. Where the data collection process was taken from 9 to 11 November 2021 to be used as a sample.

The clusters that will be formed include :

a. Cluster 1 (C1) = Loyal Customers
b. Cluster 2 (C2) = Less Loyal Customers

In this study, the initial center of the cluster or centroid was selected, namely, C1 (1, 135000), C2 (4 , 102000), and then it was possible to calculate the distance from the remaining sample data to the cluster center

$$D_i = \sqrt{(M_{1x} - C_{1x})^2 + (M_{1y} - C_{1y})^2}$$

Where D is the distance, while M is the data coordinate and C is the centroid coordinate.

Iteration-1

$$d_{1.1} = \sqrt{(4-1)^2 + (102.000 - 135.000)^2} = 33.000$$
$$d_{2.1} = \sqrt{(1-1)^2 + (21.000 - 135.000)^2} = 114.000$$
$$d_{3.1} = \sqrt{(1-1)^2 + (6.000 - 135.000)^2} = 129.000$$
$$d_{4.1} = \sqrt{(1-1)^2 + (4.000 - 135.000)^2} = 131.000$$
$$d_{5.1} = \sqrt{(1-1)^2 + (5.000 - 135.000)^2} = 130.000$$
$$d_{6.1} = \sqrt{(1-1)^2 + (8.000 - 135.000)^2} = 127.000$$
$$d_{7.1} = \sqrt{(3-1)^2 + (6.000 - 135.000)^2} = 129.000$$
$$d_{8.1} = \sqrt{(1-1)^2 + (25.500 - 135.000)^2} = 109.500$$
$$d_{9.1} = \sqrt{(5-1)^2 + (13.000 - 135.000)^2} = 122.000$$
$$d_{10.1} = \sqrt{(3-1)^2 + (3.000 - 135.000)^2} = 132.000$$
$$d_{11.1} = \sqrt{(2-1)^2 + (38.000 - 135.000)^2} = 97.000$$
$$d_{12.1} = \sqrt{(2-1)^2 + (22.000 - 135.000)^2} = 113.000$$
$$d_{13.1} = \sqrt{(1-1)^2 + (5.000 - 135.000)^2} = 130.000$$
$$d_{14.1} = \sqrt{(1-1)^2 + (2.000 - 135.000)^2} = 133.000$$
$$d_{15.1} = \sqrt{(1-1)^2 + (6.000 - 135.000)^2} = 129.000$$
$$d_{16.1} = \sqrt{(3-1)^2 + (63.000 - 135.000)^2} = 72.000$$
$$d_{17.1} = \sqrt{(1-1)^2 + (11.000 - 135.000)^2} = 124.000$$
$$d_{18.1} = \sqrt{(1-1)^2 + (135.000 - 135.000)^2} = 0$$
$$d_{19.1} = \sqrt{(3-1)^2 + (9.000 - 135.000)^2} = 126.000$$
$$d_{20.1} = \sqrt{(1-1)^2 + (10.500 - 135.000)^2} = 124.500$$
$$d_{21.1} = \sqrt{(1-1)^2 + (22.000 - 135.000)^2} = 113.000$$
$$d_{22.1} = \sqrt{(2-1)^2 + (10.000 - 135.000)^2} = 125.000$$
$$d_{23.1} = \sqrt{(3-1)^2 + (24.000 - 135.000)^2} = 111.000$$
$$d_{24.1} = \sqrt{(1-1)^2 + (9.000 - 135.000)^2} = 126.000$$

$$d_{25.1} = \sqrt{(1-1)^2 + (20.000 - 135.000)^2} = 115.000$$
$$d_{26.1} = \sqrt{(3-1)^2 + (9.000 - 135.000)^2} = 126.000$$
$$d_{27.1} = \sqrt{(1-1)^2 + (3.500 - 135.000)^2} = 131.500$$
$$d_{28.1} = \sqrt{(5-1)^2 + (5.500 - 135.000)^2} = 130.000$$
$$d_{29.1} = \sqrt{(1-1)^2 + (11.000 - 135.000)^2} = 124.000$$
$$d_{30.1} = \sqrt{(8-1)^2 + (16.000 - 135.000)^2} = 119.000$$

Do the same for calculating each point to the 2nd center

$$d_{1.1} = \sqrt{(4-4)^2 + (102.000 - 102.000)^2} = 0$$
$$d_{2.1} = \sqrt{(1-4)^2 + (21.000 - 102.000)^2} = 81.000$$
$$d_{3.1} = \sqrt{(1-4)^2 + (6.000 - 102.000)^2} = 96.000$$
$$d_{4.1} = \sqrt{(1-4)^2 + (4.000 - 102.000)^2} = 98.000$$
$$d_{5.1} = \sqrt{(1-4)^2 + (5.000 - 102.000)^2} = 97.000$$
$$d_{6.1} = \sqrt{(1-4)^2 + (8.000 - 102.000)^2} = 94.000$$
$$d_{7.1} = \sqrt{(3-4)^2 + (6.000 - 102.000)^2} = 96.000$$
$$d_{8.1} = \sqrt{(1-4)^2 + (25.500 - 102.000)^2} = 76.500$$
$$d_{9.1} = \sqrt{(5-4)^2 + (13.000 - 102.000)^2} = 89.000$$
$$d_{10.1} = \sqrt{(3-4)^2 + (3.000 - 102.000)^2} = 99.000$$
$$d_{11.1} = \sqrt{(2-4)^2 + (38.000 - 102.000)^2} = 64.000$$
$$d_{12.1} = \sqrt{(2-4)^2 + (22.000 - 102.000)^2} = 80.000$$
$$d_{13.1} = \sqrt{(1-4)^2 + (5.000 - 102.000)^2} = 97.000$$
$$d_{14.1} = \sqrt{(1-4)^2 + (2.000 - 102.000)^2} = 100.000$$
$$d_{15.1} = \sqrt{(1-4)^2 + (6.000 - 102.000)^2} = 96.000$$
$$d_{16.1} = \sqrt{(3-4)^2 + (63.000 - 102.000)^2} = 39.000$$
$$d_{17.1} = \sqrt{(1-4)^2 + (11.000 - 102.000)^2} = 91.000$$
$$d_{18.1} = \sqrt{(1-4)^2 + (135.000 - 102.000)^2} = 33.000$$
$$d_{19.1} = \sqrt{(3-4)^2 + (9.000 - 102.000)^2} = 93.000$$
$$d_{20.1} = \sqrt{(1-4)^2 + (10.500 - 102.000)^2} = 91.500$$
$$d_{21.1} = \sqrt{(1-4)^2 + (22.000 - 102.000)^2} = 80.000$$
$$d_{22.1} = \sqrt{(2-4)^2 + (10.000 - 102.000)^2} = 92.000$$
$$d_{23.1} = \sqrt{(3-4)^2 + (24.000 - 102.000)^2} = 78.000$$
$$d_{24.1} = \sqrt{(1-4)^2 + (9.000 - 102.000)^2} = 93.000$$
$$d_{25.1} = \sqrt{(1-4)^2 + (20.000 - 102.000)^2} = 82.000$$
$$d_{26.1} = \sqrt{(3-4)^2 + (9.000 - 102.000)^2} = 93.000$$
$$d_{27.1} = \sqrt{(1-4)^2 + (3.-102.000)^2} = 98.500$$
$$d_{28.1} = \sqrt{(5-4)^2 + (5.500 - 102.000)^2} = 97.000$$
$$d_{29.1} = \sqrt{(1-4)^2 + (11.000 - 102.000)^2} = 91.000$$
$$d_{30.1} = \sqrt{(8-4)^2 + (16.000 - 102.000)^2} = 86.000$$

Table 1First Iteration Object Distance

Data i	C1	C2	Clusters
1	33,000	0	2
2	114,000	81,000	2
3	129000	96000	2
4	130,000	98.00	2
5	130000	97.00	2
6	127000	94000	2
7	129000	96000	2
8	109500	76500	2
9	122000	89000	2
10	132000	99000	2
11	97000	64000	2
12	113000	80000	2
13	130000	97000	2
14	133000	10000	2
15	129000	96000	2
16	72000	39000	2
17	124000	91000	2
18	0	33000	1
19	126000	93000	2
20	124500	91500	2
21	113000	80000	2
22	125000	92000	2
23	111000	78000	2
24	126000	93000	2
25	115000	82000	2
26	126000	93000	2
27	131500	98500	2
28	130000	97000	2
29	124000	91000	2
30	119000	86000	2

After the distance of each object is known, the data is immediately allocated to *the cluster* that has the closest *centroid* as stated in. Table 3. Then determine the new *centroid value* with the formula listed above. The results of this process are as follows:

Table 2First Iteration Data Allocation

Data i	Clusters 1		Clusters 2	
	X	Y	X	Y
1	0	0	4	102,000
2	0	0	1	21,000
3	0	0	1	6,000
4	0	0	1	4,000
5	0	0	1	5,000
6	0	0	1	8,000

7	0	0	3	6,000
8	0	0	1	25.5000
9	0	0	5	13,000
10	0	0	3	3,000
11	0	0	2	38,000
12	0	0	2	22,000
13	0	0	1	5,000
14	0	0	1	2,000
15	0	0	1	6,000
16	0	0	3	63,000
17	0	0	1	11,000
18	1	135,000	0	0
19	0	0	3	9,000
20	0	0	1	10,500
21	0	0	1	22,000
22	0	0	2	10,000
23	0	0	3	24,000
24	0	0	1	9,000
25	0	0	1	20,000
26	0	0	3	9,000
27	0	0	1	3,500
28	0	0	5	5,000
29	0	0	1	11,000
30	0	0	8	16,000
Average	1	135,000	2,1	16,879

After the iteration-1 process, the new centroid center positions will be obtained, namely C1 (1, 135000) and C2 (2,1, 16879). The search will continue to iteration-2 with the same process as iteration-1. Then the following results will be obtained:

Table 3Second Iteration Object Distance

Data i	C1	C2	Clusters (Old)	Clusters (New)
1	33000	85121	2	1
2	114000	4121	2	2
3	129000	10879	2	2
4	131000	12879	2	2
5	130000	11879	2	2
6	127000	8879	2	2
7	129000	10879	2	2
8	109500	8621	2	2
9	122000	3879,001	2	2
10	132000	13879	2	2
11	97000	21121	2	2
12	113000	5121	2	2
13	130000	11879	2	2
14	133000	14879	2	2
15	129000	10879	2	2
16	72000	46121	2	2
17	124000	5879	2	2
18	0	118121	1	1
19	126000	7879	2	2
20	1245000	6279	2	2
21	113000	5121	2	2
22	125000	6879	2	2
23	111000	7121	2	2
24	126000	7879	2	2
25	115000	3121	2	2
26	126000	7879	2	2
27	1315000	13379	2	2
28	130000	11879	2	2
29	124000	5879	2	2
30	119000	879,0198	2	2

After the distance of each object is known, the data is immediately allocated to *the cluster* that has the closest *centroid* as shown in Table 5. Then determine the new *centroid value* with the formula listed above. The results of this process are as follows:

Table 4Allocation of Second Iteration Data

Data i	Clusters 1		Clusters 2	
	X	Y	X	Y
1	4	102,000	0	0
2	0	0	1	21,000
3	0	0	1	6000
4	0	0	1	4000
5	0	0	1	5000
6	0	0	1	8000
7	0	0	3	6000
8	0	0	1	25500
9	0	0	3	169,000
10	0	0	3	3000
11	0	0	2	38000
12	0	0	2	22000
13	0	0	1	5000
14	0	0	1	2000
15	0	0	1	6000
16	0	0	3	63000
17	0	0	1	11000
18	1	135000	0	0
19	0	0	3	9000
20	0	0	1	10500
21	0	0	1	22000
22	0	0	2	10000
23	0	0	3	24000
24	0	0	1	9000
25	0	0	1	20000
26	0	0	3	9000
27	0	0	1	3500
28	0	0	5	5000
29	0	0	1	9000
30	0	0	8	16000
Average	2.5	118,500	2.07	13,839

After the second iteration process, the new centroid center positions will be obtained, namely C1 (2,5 , 118500) and C2 (2,07, 13839). The search will continue to iteration-3 with the same process as the previous iteration. Then the following results will be obtained:

Table 5Third Iteration Object Distance

Data i	C1	C2	Clusters (Old)	Clusters (New)
1	16500	88161	1	1
2	97500	7161	2	2
3	112500	7839	2	2
4	114500	9839	2	2
5	113500	8839	2	2
6	110500	5839	2	2
7	112500	7839	2	2
8	93000	11661	2	2
9	105500	839,0051	2	2
10	115500	10839	2	2
11	80500	24161	2	2
12	96500	8161	2	2
13	113500	8839	2	2
14	116500	11839	2	2
15	112500	7839	2	2
16	55500	49161	2	2
17	107500	2839	2	2
18	116500	121161	1	1
19	109500	4839	2	2
20	108000	3339	2	2
21	96500	8161	2	2
22	108500	2829	2	2
23	94500	10161	2	2
24	109500	4839	2	2
25	98500	6161	2	2
26	109500	4839	2	2
27	1315000	115000	2	2
28	13500	8839	2	2
29	124000	5879	2	2
30	119000	879,0198	2	2

After the distance of each object is known, the data is immediately allocated to *the cluster* that has the closest *centroid* as shown in Table 7. Then determine the new *centroid value* with the formula listed above. The results of this process are as follows:

Table 6Third Iteration Data Allocation

Data i	Clusters 1		Clusters 2	
	X	Y	X	Y
1	4	102,000	0	0
2	0	0	1	21,000
3	0	0	1	6000
4	0	0	1	4000
5	0	0	1	5000
6	0	0	1	8000
7	0	0	3	6000
8	0	0	1	25500
9	0	0	3	169,000
10	0	0	3	3000
11	0	0	2	38000
12	0	0	2	22000
13	0	0	1	5000
14	0	0	1	2000
15	0	0	1	6000
16	0	0	3	63000
17	0	0	1	11000
18	1	135000	0	0
19	0	0	3	9000
20	0	0	1	10500
21	0	0	1	22000
22	0	0	2	10000
23	0	0	3	24000
24	0	0	1	9000
25	0	0	1	20000
26	0	0	3	9000
27	0	0	1	3500
28	0	0	5	5000
29	0	0	1	9000
30	0	0	8	16000
Average	2.5	118,500	2.07	13,839

The search is stopped because the number of members from this iteration and the previous iteration has not changed. With this, it is known that each cluster of loyal customers and less loyal customers is based on the number of transactions and total transactions.

Table 7*Clustering* Results

Iteration	Centroid value used				Cluster Member	
	C1		C2		1	2
	X	Y	X	Y		
1	1	135,000	4	102,000	29	1
2	1	135,000	2,1	16,879	28	2
3	2.5	118,500	2.07	13,839	28	2

In the calculations that have been done above, it can be concluded that of the 30 transaction data provided it produces two groups (clusters) with the condition that the first cluster (Loyal Customers) produces 2 members with customer id (001 & 018) and the second cluster (Less Loyal Customers) generate 28 members with customer id (002 ,003 ,004 , 005 , 006 , 007 ,008 , 009 , 010 , 011 , 012 ,013 ,014 ,015 , 16 , 017 , 019 , 020 , 021 , 022 , 023, 024 025 ,026 ,027 028 ,029 030). And it can be stated that customers with clusters (Less Loyal Customers) are clusters that have the most members in clustering based on the number of transactions and total transactions.

3.2 Testing the K-Means Process on the System

1. *Customer clustering* processing is carried out on the *admin page* where the data used for K-means processing is obtained from transactions that have been made by *customers* on the Rani Swalayan 2 *website* , these data will continue to grow as more transactions are made by *customers* . Like the following picture:

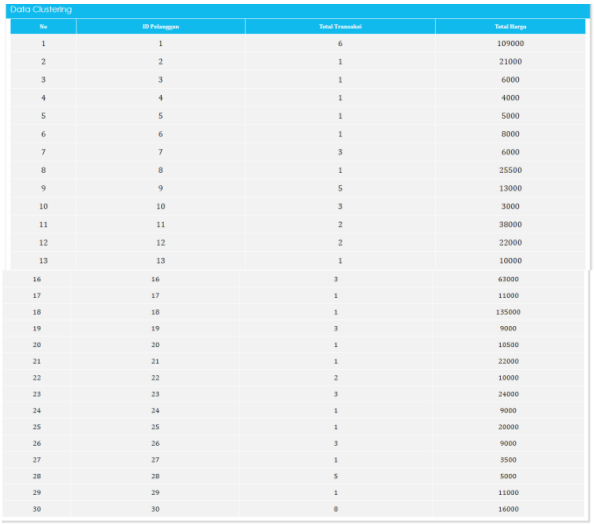

Figure 2 Data *Clustering*

2. After the *customer clustering data* is available, then carry out the clustering process by determining the center point (*centroid*) which is determined from the existing transaction data which is taken randomly *then* clicking the "Process" button, as shown in the following image:

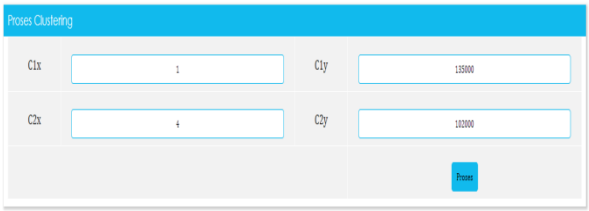

Figure 3 Clustering Process

3. After the *clustering process* is carried out, the system will display the *clustering results* obtained consisting of C1 and C2 results, data grouping results and detailed K-means calculations performed, as shown below:

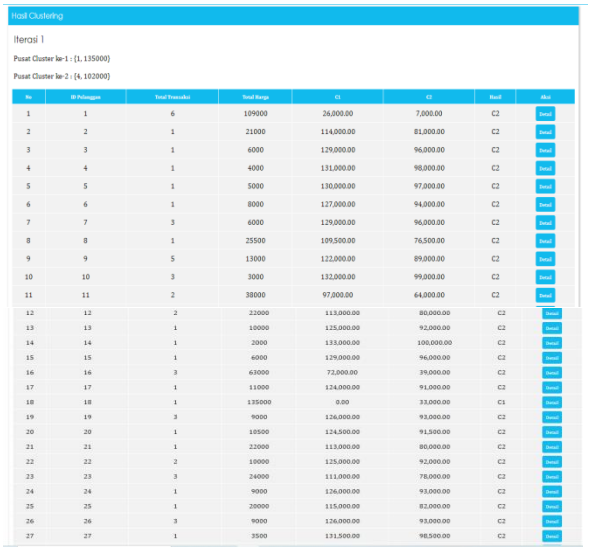

Figure 4 Clustering Results

4. Conclusion

Customer Relationship Management (CRM) application using the web-based K-Means algorithm using 30 data samples. has been able to optimize customer loyalty. This is evidenced by the many transactions made at Rani Swalayan 2.

References

[1] Hia, HH, Saragih, NF, & Larosa, FGN (2019). *Application of CRM in the Gunungsitoli City Tax Counseling and Consultation Service Office Application (KP2KP). 3(2), 97–106. https://doi.org/10.31227/osf.io/8bmc4*

[2] Irawan, Y. (2019). No TitleΕΛΕΝΗ. Αγαη, 8(5), 55.

[3] Kartaman, AT, & Winangun, AK (2019). CRM (Customer Relationship Management) Based E-Business Design Case Study: Shoe Washing Service Company S-Neat-Kers. In IDEC National Seminar and Conference. .

[4] Nooraeni, R. (2015). The Cluster Method Uses a Combination of the K-Prototype Cluster Algorithm and the Genetic Algorithm for Mixed Type Data. Journal of Statistical Applications & Statistical Computing, 7(2), 17-17. *N* . https://doi.org/10.34123/jurnalasks.v7i2.23

[5] Alvisan, F, K. (2021). *Minimarket Clustering to Determine the Number of Purchase Needs Using the K-Means Method* . NOE Journal. https://doi.org/10.29407/noe.v4i2.16784

[6] Joyendri, A. (2017). Customer Relationship Management Strategy To Increase Customer Loyalty And Sales Volume Using The K-Means Clustering Technique. Telematics: Journal of Informatics and Information Technology, 14(2), 75-82.. https://doi.org/10.31315/telematika.v14i2.2094

[7] Puteri, ET, Kusnanto, G., & Thomas, CJ (2020). Application of K-Means Clustering for Customer Segmentation in the Customer Relationship Management System at Pt. Unichem Indonesian Temple..

[8] Abdulloh, R. (2018). 7 in 1 Web Programming For Beginners. Elex Media Komputindo.

[9] Kholil, I. (2017). Web-Based Customer Relationship Management (CRM) To Improve Online Store Competitiveness. Journal of Pilar Nusa Mandiri, 13(1), 43-48. https://doi.org/10.33480/pilar.v13i1.145

[10] Mulyodiputro, MD (2018). Pharmacy Information System Database Design Using MySQL at Cemara Pharmacy. ScienceTech Innovation Journal, 1(1), 16-19. https://doi.org/10.37824/sij.v1i1.2018.17