Customer Loyalty Segmentation on Point of Sale System Using Recency-Frequency-Monetary (RFM) and K-Means

Bayu Rizki¹, Nava Gia Ginasta², Muh Akbar Tamrin³, Ali Rahman⁴
¹,²,³Department of Informatics Engineering, Politeknik Pos Indonesia, Indonesia
⁴Department of Magister Information System, STMIK Likmi, Indonesia

Article Info

Article history:
Received February 19, 2020
Revised May 07, 2020
Accepted September 02, 2020
Published December 30, 2020

Keywords:
K-Means
Point of sale
RFM

Corresponding Author:
Bayu Rizki
Department of Informatics Engineering,
Politeknik Pos Indonesia
Jl. Sariasih. No.54, Bandung, Indonesia
Email: bayurizki135@gmail.com

ABSTRACT

It is no doubt that the development of the business world has been progressive. Point of sale is one of the many systems used as a means of payment in various existing businesses, especially in heterogeneous markets. The activity of transactions between Point of Sale Systems and Customers occur in the business world. Keep in mind also that one of the main factors of business success, is from customers. There is the need of an attractive strategy and certainly it will be to increase the income and assets of a business. To know that, this research will explore the behavior of customer which is based marketing, through RFM Method (Recency, Frequency and Monetary). The case of this study is in Goldfinger Store. It will do segmentation and also use data mining technique to do clustering by using K-Means with result of loyal and potential customer. The results of segmentation using RFM (Recency, Frequency, Monetary) and K-Means methods have produced multiple clusters by dividing them into groups.

1. INTRODUCTION

The development of technology without stopping, indicating the era of platform providers more rapidly. Currently, technology has become the company’s primary requirement in carrying out its work[1]. Sobat Technology which is one software company located in Bandung, supporting as well as assisting business integration into Information Technology system. The demand for innovation and creativity, get some input from the client, call it Goldfinger Store. Request the application of technology-based business strategies, to create a system that supports the marketing, of data Systems Point of Sale (POS) through a web service and start work functionally and adequately[2]. Competition for the business world drives some companies to manage customers well. In general, customers who can benefit a high-quality customer-owned company[3]. The need for a strategy that is capable of maintaining customer quality as well as a way to increase the company’s asset revenue. Seeing the problems that occur Goldfinger Store has not been able to determine customer data information that can provide more benefits for the company, what happens is only limited to the library of transaction data used in the calculation of profit only. In This is the company’s difficulty in determining the marketing strategy because of the ignorance of customer characteristics. In business, marketing is a critical part of the direct sales industry[4]. As for the aspects not only in the Databank marketing apply retaining customer loyalty[5]. By understanding the aspects of customers, it is necessary to collect market-related information[6].

In this study, the author will determine the potential and loyal customers, to the progress of the company by maintaining the relationship between customers. The term is trendy in sales and marketing, like CRM (Customer Relationship Management). Implement CRM (Customer Relationship Management) to perform the correct marketing strategy so that it can bring benefits to the company[7]. The Actually, the implementation of CRM is a way that has many applications to do customer segmentation. In a sense segmentation is the process of taking advantage of opportunities by dividing the market into several segments[8]. Customer segmentation and various customer results are useful for an explanation of a company’s
marketing strategy[9]. Moreover, with the cementation of capable subscribers, the company will be able to provide customer groups to maximise profits[10]. That way will be seen comparison through actual POS system sales data from the company[11].

For purposes, research uses POS Data (Points Of Sales) to extract purchasing behaviour for CRM (Customer Relation Management)[12]. On the other hand, it takes a useful tool for customer segmentation in the selection of attributes to get better results. RFM (Recency, Frequency and Monetary) models are one of the most appropriate methods for customer description descriptions, which consequently will affect increasing sales[13]. In recent years, the RFM model has been referring to customer behavioural analysis, based on the level of eligibility, transaction frequency, and monetary background[14]. However, of course it is essential to add Data Mining Techniques in a customer segmentation clustering. Mainly, data mining is a methodology of extracting patterns from data[15]. K-Means algorithm is one of the most common clustering techniques in use because it can group customers into different categories[16]. As is known, the K-Means Algorithm can classify data by finding the closest distance to the Cluster centre point[17]. In its analysis of data on just one the next, then the system has not yet found a pattern of optimally. K-Means will divide all the customers into the appropriate cluster group. Therefore, examination of the characteristics of each cluster will be beneficial to be able to determine and retain profitable and loyal customers. Then it will efficiently develop regarding marketing strategy for each customer cluster[18]. Customers will segment with a k-means algorithm based on customer payment information to measure customer’s potential level[19].

2. METHOD

Stages The research methodology that will be used is problem identification, data collection, data analysis, data processing with the RFM method as determining indicators and the K-Means Method as the final grouping and objectives to be achieved. The stages of research undertaken in conducting this research are depicted in Figure 1.
A. The RFM Model
The RFM model is a useful segmentation method for market segmentation through effective analysis[18]. RFM analysis of the three parameters is important for the Recency, F for Frequency, and M for Monetary[9]. In theory RFM analysis is assumed (P). The description of three basic characteristics:

\[ P = \langle R, F, M \rangle \]  

R - the last purchase reviewer - is calculated as the difference between the current date and the date of the last purchase given in the day decision
F - the frequency of purchase - the number of purchases during the observation period.
M - monetary - the total amount of capital[13].
More specifically, Recency refers to the time interval concerning previous transactions. Frequency represents no. of transactions made by the customer within a specified period. So the next consumer purchase can make predictions. Monetary, denotes a certain amount that is exhausted by individuals[14]. The RFM model can efficiently achieve customer relationship management. This model is one of the essential means to measure the profitability of customer value[10].

![Figure 2. The RFM Model](image)

B. K-Means
K-means is the most famous algorithm in clustering techniques and was first introduced by Macqueen[20]. K-means is also one of the most prominent grouping techniques in science and technology[21]. Grouping data that has the same characteristics in one cluster or grouping data groups that have characteristics different from other groups or groups so that data in one cluster or group has a small level of variation[22]. This well-known algorithm of K-means can determine the number of groups at the beginning and define a set of K centroids. The K-means K-means algorithm will in detail introduce three aspects: distance measuring, cluster centre, and algorithmic processes[23]. The abundance of K-means, in application to various data types because of its simplicity. The following steps in the K-means algorithm as follows:
Customer Loyality Segmentation on Point of Sale System Using Recency-Frequency-Monetary (RFM) and K-Means

(Bayu Rizki*, Nava Gia Ginasta, Muh Akbar Tamrin, Ali Rahman)

Figure 3. Flow chart of K-means Algorithm

Step 1: Initialize the number of clusters and the average value of the cluster.
Step 2: Calculate the distance of the image intensity of the tool of all clusters.
Step 3: Cluster and label the image based on calculated distance.
Step 4: Calculate the new average value for each cluster.
Step 5: Repeat steps 2, 3 and four until no more change in the calculated mean value.

K-means formula for calculating distance using Euclidean

\[ d(X_j, C_j) = \sqrt{\sum_{i=1}^{n} (X_{ij} - C_{ij})^2} \]  \hspace{1cm} (2)

\[ D(i,j) = \sqrt{(X_{i1} - X_{j1})^2 + (X_{i2} - X_{j2})^2 + \cdots + (X_{ik} - X_{jk})^2} \]  \hspace{1cm} (3)

d = distance, j = amount of data, c = centroid, x = data, c = centroid D (i, j) is the data distance between i and the centre cluster j. X \( k \) is the data for me attribute data to k. X \( k \) j is the centre point to j at attribute to k. Recalculate the cluster centre with the current cluster membership. The cluster centre is the average of all data/objects on a particular cluster [17].

3. RESULTS AND DISCUSSION

3.1. Data Extraction using RFM

The data used as the case example is the client transaction data from Goldfinger Store period December 2016-November 2017.

1. Determining the Point Recency, then the first set of customer segmentation rules as follows:
   Segmentation rules (number of days 1 period-end date period)
   1 December 2016 - November 30, 2017 = 364 days
   \[ \frac{364 \text{ days}}{2} = 182 \text{ days} \]
   Recency Point:
   a. Recency \( \leq 182 \) days then include customer with point 3 scale
   b. 183 days \( \leq \) Recency \( \geq 364 \) days later including customer with the size of point 2
   c. Recency \( \geq 365 \) days then including customers with a point scale of 1
2. Determine point Frequency
Calculation of frequency as follows:
Frequency = \( \frac{\text{number of frequency}}{\text{quantity of goods}} \)
Frequency = \( \frac{106}{16} = 6 \), because using 3 segmentation then this average number should be divided by 3, then this means:
\[\frac{6}{3} = 2\]
\[2 \times 2 = 4\]
Frequency 2 is a non-potential counter (Rare), with a point scale of 1
Frequency 4 counters potential (Middle), with a point scale of 2
Frequency > 5 is a possible counter (Frequent) with a point scale of 3.

3. Determining the point of Monetary
Calculation as follows,
Monetary = \( \frac{\text{Total Monetary}}{\text{Amount Of Customers}} \)
\[= \frac{27,542,000}{71 \text{ Customers}}\]
\[= 387,000 (387.915)\]
\[= \frac{387,000}{3}\]
\[= 129,000\]
\[= 129,000 \times 2 = 258,000\]

Monetary
a. Monetary ≤ 129,000 then including customer with scale point 1.
b. 129001 < Monetary ≥ 258,000 then including customer with scale of point 2.
c. Monetary > 258,000 then including customer with scale of point 3.

3.2. Weighting Using RFM
The results of the weighting process using RFM Model.

<table>
<thead>
<tr>
<th>Customer’s name</th>
<th>R</th>
<th>F</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efendi</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Taufik Hidayat</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Faadil Saadillah</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Muhammad Ilham iqra</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Anwar</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Putra Wibowo</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Rafli Muhammad</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Arfan</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Miftahul khaeri</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Riadil asad</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Faiz Muhajir</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Amal Bakti</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Kurnia Saputra</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Kevin</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

3.3. The Cluster Process Using K-Means
The Results from the Cluster Process Using K-Means after weighting using RFM Model.

<table>
<thead>
<tr>
<th>Customer Name's</th>
<th>R</th>
<th>F</th>
<th>M</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arfan</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>C1</td>
</tr>
<tr>
<td>Miftahul khaeri</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>C1</td>
</tr>
<tr>
<td>Riadil asad</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>C1</td>
</tr>
</tbody>
</table>

166,766
### Table 3. Results of the K-Means Method Component Grouping (Medium)

<table>
<thead>
<tr>
<th>Customer Name's</th>
<th>R</th>
<th>F</th>
<th>M</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taufik Hidayat</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>C2</td>
</tr>
<tr>
<td>Putra Wibowo</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>C2</td>
</tr>
<tr>
<td>Rafli Muhammad</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>C2</td>
</tr>
<tr>
<td>Faiz Muhajir</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>C2</td>
</tr>
<tr>
<td>Amal Bakri</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>C2</td>
</tr>
<tr>
<td>Kurnia Saputra</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>C2</td>
</tr>
<tr>
<td>Kevin</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>C2</td>
</tr>
</tbody>
</table>

2,167

### Table 4. Results of the K-Means Method Component Grouping (Height)

<table>
<thead>
<tr>
<th>Customer Name's</th>
<th>R</th>
<th>F</th>
<th>M</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eleni</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>C3</td>
</tr>
<tr>
<td>Faadil Saadillah</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>C3</td>
</tr>
<tr>
<td>Muhammad Ilham iqra</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>C3</td>
</tr>
<tr>
<td>Anwar</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>C3</td>
</tr>
</tbody>
</table>

2,667

### 3.4. RFM Results With K-Means in Point of Sale Applications

The results of the implementation of RFM (Current, Frequency, Monetary) and grouping using the k-means algorithm in the point of sale application. In Figure 4 it is divided into three centroids, and the grouping is based on the smallest distance. RFM implementation for weighting before grouping uses k-means method to determine customer data that has the highest, medium, and low loyalty.

![Figure 4. Result in point of sale application](image)

### 4. CONCLUSION

RFM method is the initial method used in this research. Recency, Frequency, Monetary methods can help the company in knowing customer loyalty. And in particular, can help and show the optimal cluster quality. K-Means method is a clustering method used in the final stages of this study. K-means method is very helpful in customer clustering as the final solution in this research.

### 5. REFERENCES


Customer Loyalty Segmentation On Point Of Sale System Using RFM And K-Means

(Bayu Rizki¹, Nava Gia Ginasta², Muh Akbar Tamrin³, Ali Rahman⁴)


