

GURILEM : A Novel Design of Customer Rating Model using K-Means and RFM

Rolly Maulana Awangga¹⁾, Syafrial Fachri Pane²⁾, Diana Asri Wijayanti³⁾

Applied Bachelor Program of Informatics Engineering
Politeknik Pos Indonesia, Jl. Sariasih No 54 Bandung, 40151, Indonesia
Email: awangga@poltekpos.ac.id¹⁾, syafrial.fachri@poltekpos.ac.id²⁾,
dianaasriwijayanti@if.poltekpos.ac.id³⁾

Abstract

A rating system or reviews are generally used to assist in making decisions. Rating systems widely used as a technique in recommending items to customers, one of which used in recommending resorts to be used by customers. However, the credibility of the rating looks vague because the rating could only represent some points of service. So that customer preference with each other is very different. Personalized recommendation systems offer more personalized advice, precisely knowing the preferences or tastes of the customers. Especially for customers who have a transaction history or reservation as at their resorts provide useful information used by managers to design a recommendation model for their customers. In this study aims to create a model of resort recommendations based on a rating of frequency. This frequency is the number of the resort used by the customer within the specified time frame. With the frequency can represent the preferences of customers. The RFM (Recency, Frequency, and Monetary) method is used to measure the reservation frequency value of the customer. The K-Means method is used to categorize customer data with its frequency and classify the type of resort. Recommendation resort to the customer based on the dominant use in one of the resort types. The type of resort recommended for customers base on the similarity between the type of resort used and other types of resorts.

Keywords: customer preferences, rating, recommendation, RFM, K-Means.

1. INTRODUCTION

A recommendation system is a technology that seeks to find products and offers or services that users may be interested [1], [2] or a particular type of information system that helps decision-makers choose items that fit their preferences and interests [3]. Technology has a vital role in helping the

Commented [A1]: Revision of point number 1

company's performance [4], [5]. The Recommender system has been widely used in many online sites to help customers address information overload and make their purchasing decisions [6]. The recommendation system offers personalized suggestions by analyzing user preferences [2]. Preferences can be found from the custom done [7]. The personalized search aims to help users identify desired products or services based on their personal preferences [1] specifically for travelers providing user history information [8]. In the recommended application, there is a need to recommend a user-selectable package [9]. Sophisticated recommendations result in higher customer satisfaction [10]. Customers always want to get maximum satisfaction from the product or service they buy [11]. One example of techniques in a recommendation system is the use of a rating that is typically a user-assigned star of an item. In that case, the user assesses an item he once used. For example, in the assessment of a hotel, in some hotels that have used a rating system such as the star rating that users typically give stars on a scale of 1-5 [12]. Another example would be to recommend a movie where the value is 1 "Hated It"; the value is 2 "Did not Like"; the value is 3 "Liked It"; 4 "Really Liked It"; and 5 "Loved It" [10]. The star rating assessment system usually assesses in its entirety, so the assessment is more specific in some points such as the service, and its cleanliness is poorly used and noticed [13]. Because it could be on the other hand, the user judges poorly in one thing but judges both in some ways, so the user prefers good judgment [14] and uses fewer reviews. Today's assessment system not only allows consumers to read and write reviews about products but also to check the credibility of reviews [15]. A widely used recommendation system usually based on similarities between active users and other users [16]. However, unlike the personalized recommendation system where recommendations are given based on user preferences [1].

The use of rating with star-scale still needs to be evaluated for its credibility. The rating that only uses a star scale can be unfair or data engineering as much happens at this time. The preferences of an individual cannot know if by looking at similarities between users with each other, which is generally based on the background of the user.

2. RELATED WORKS

Procedure recommendations are given based on known user preferences through user ratings continuously as recommendations arise. With so can know the preferences of the user [17]. On film recommendations to users by studying ratings provided by users can give users a sense of preference regarding other types of movies, making it easy to use to make movie recommendations to users [18] searching for one user's preference equations with other users is done on the cluster. Each user has different preferences [17], [19]. Customer data from social media networks are used to find user preferences. One technique for recommendations based on exact visit times is used rather than other measurement techniques [20]. The

proposed recommendation approach personalized incorporates three social factors; one of them is the user's interest. User interest is another important factor for influencing user decision making processes [21]. Personalized an excellent recommendation system service requires the use of various types of data needed, for example, that is such as product purchase history by users [16]. For the travelers, the recommendation system is an efficient way in search of a tourist attraction that matches their preferences [22] because it can overcome the problem of information overload to determining decision [23]. Such as a collection of images that exist on social media can provide information on the level of visits on the tourist attraction [24]. The proposed tourism recommendation system allows users to obtain sufficient information at their request and enjoy a fun tour [25]. Travel tour recommendations have constraints such as interest preferences and the needs of different tourism tours, and every tourist has uniqueness [26]. The recommendation system offers personalized suggestions by analyzing user preferences. Choosing a suitable hotel can be very important for an enjoyable trip. In any tourist travel, the selection of hotels must be the primary thing [27]. The selection of features that interest users can indicate the user's interest profile [6].

The RFM method is used to analyze and identify customer behavior [28], [29]. Moreover, the RFM used to know the segmentation of customer data [30]. The purpose of RFM is to predict future customer behavior. Therefore, it is necessary to translate customer behavior in 'numbers' so that it can use all the time [31]. The identify the customer profile with examines the customer transaction patterns; this is required to fill the recency variable used to measure the time-based visit scores. The frequency variable is the number of visits during the specified period. The monetary variable is the total cost incurred by the customer to pay for the service [29]. The identifier customer profile, the company can build a better market segmentation strategy and improve customer satisfaction. The data used is from a customer purchase database that contains customer transactional data [28].

K-means is also used to create the procedure of autocomplete feature to search for coal term. The K-Means algorithm is used to classify data from coal terms based on words and characters. With this grouping can make it easy to make predictions or suggestions in a term search [5]. In the recommendation system, use of such as data from demographic groups based on user descriptions, for example, interest or preference. Over time the number of preference groups will continue to grow due to increasing users. The K-means method helps in categorizing customer preferences so more accessible to identify and be used [32]. The grouping of user data consists of categories specified by the admin. From the clustering results using K-Means, some clusters include multiple users. The value of centroid use as a recommendation of a destination for users in the group recommended travel destinations rank by category with the highest to lowest [33]. The K-Means method too used to partition the user-item-context interactions. All users in

the cluster considered as like-minded users. When new users express their preferences, the system compares the distance between existing users and groups and adds users to the appropriate group [34].

3. ORIGINALITY

In this research, it aims to make resort recommendations for customers based on reservation history data with the case study on Agrotourism N8 owned by PT. Perkebunan Nusantara VIII. Agrotourism N8 is a lodging service provider. Agrotourism N8 produces daily data assets of customer order history that are useful for improving the quality of service, one of which is for the recommendation system provided to customers. This recommendation system will help officers to provide resort recommendations to customers by knowing customer preferences. So that with this recommendation system, Agrotourism N8 has the advantage of managing its customers, besides providing satisfaction to customers.

The purpose of this research intends to create a model of resort recommendations design with a rating based on the frequency of customers in resort use and at the same time as a customer rating. Identification of customer preferences is essential to be used to provide appropriate recommendations that can generate customer satisfaction. One way is to know customer preferences in real terms. Identification of preferences from customers using frequency variables in this study can see in real terms. The form of the customer's preference reality is the use of a service in a repetitive. The method of frequency as a rating is more accurate and looks real so that the credibility of the assessment of a frequency is visible. Customer preferences are essential to know because its indicated for recommendations that generate satisfaction for customers and to provide a positive impact on the company, for example, to generate customer loyalty. The model design for this recommendation procedure is expected to used as a means to help provide choice to customers. The items offered will be tailored to their preference. The method used is RFM and K-Means. Data processing divided into three stages. The first stage is to get the frequency value from the customer's reservation history data. The second is uses data from the RFM method results along with customer reservation information in the form of selected resorts. The third stage is the grouping of types of resorts. Type of resort recommendations gave to customers based on their preferences known through the reservation history.

4. THE METHOD

4.1 RFM (Recency, Frequency, Monetary)

The RFM (Recency, Frequency, Monetary) method is used to analyze and identify customer behavior based on three variables. Recency variable used to measure the time of the last purchase period and the latest purchase by the customer based on date, day, or month. The Frequency variable is measuring the number of purchases. The monetary variable is to measure the

Commented [A2]: Revision of point number 2

Commented [A3]: Revision of point number 2

number of costs incurred by the customer based on a specific time. The RFM method helps companies to know the behavior of customers in transacting and helps to set up appropriate marketing strategies [28], [29]. The RFM method, with its attributes, is an appropriate method for performing and knowing the segmentation of customer data [30].

4.2 K-Means

The K-Means algorithm is one of the most commonly used grouping methods because of its simplicity, flexibility, and computational efficiency, especially in managing extensive data [35]. So the K-Means algorithm is easier to use to describe the characteristics of each group or cluster. K-Means is a simple algorithm and has a reasonably high accuracy according to the size of the object. Also, the K-Means algorithm is not affected by the order of objects [33].

K-Means iteratively calculates the mean value of K to set the purpose into the closest cluster based on distance [35]. The numerical data in the form of numbers which can only process. Data other than numerical can also be treated but must be changed by represented by the code to facilitate the calculation of distance or similarity characteristics on each object [33].

The formula for calculating distances on the K- Means method uses the Euclidean Distance formula as follows :

$$d(X_j, C_j) = \sqrt{\sum_{j=1}^n (X_j - C_j)^2} \quad (1)$$

On the formula 1 where d is the distance, j represents the amount of data, C is centroid, X is data, and C is centroid [5]. From the calculation of the midpoint of the data will be found and as a cluster point called the centroid, then other data is placed in the nearest cluster. The centroid can found when all data is in the closest group by calculating the average data in the cluster. First determine the number of k, as the number of clusters to be formed C_j; j = 1; 2; 3; k. Second, set the center point of the cluster k at random. The third calculates the distance of each data to the center point of the cluster (centroid). Fourth Input each data into the nearest centroid, then the fifth species the new centroid point by calculating the average value of data that is in the same centroid point and the last six to repeat step 3 if the position of new and old centroid points are different [33].

5. EXPERIMENT AND ANALYSIS

5.1 Experiment

5.1.1 RFM

The application of the RFM method in this study is to process customer reservation history data that focuses on knowing the frequency

Commented [A4]: Revision of point number 2

Commented [A5]: Revision of point number 2

value of customer reservations for resort types. The obtained frequency value used as a rating on the resort. The variable F (frequency) is used as the amount of rent or resort usage by the customer for one year.

Table 1. Example of data reservation history

Name or Institution	Resort Type Booked	Check-In Date
RC001	Kidang Kencana	02/1/2016
RC002	Kidang Kencana	31/1/2016
RC003	Cottage	15/1/2016
RC004	Villa Patenggang	1/1/2016 - 2/1/2016
RC005	Kidang Kencana	30/1/2016
RC062	Ciung Wanara	26/1/2016 & 23/6/2016
RC062	Dungus Wangi	26/1/2016 & 23/6/2016
RC062	Kidang Kencana	23/6/2016

Table 1 is an example of a data reservation history. Reservation history data is converted into a customer-ordered RFM model to get frequency value. For information, the Name or Institution column changed to code (id). Customer reservation history data based on data collection in one year (one period). Frequency value viewed from Check-in Date. Customers who choose one type of resort in one day valued at 1 for the frequency. The range of rating values generated from the calculation results is free because the rating to used as a recommendation is the highest rating produced by each customer. The higher the rating value of the customer is a priority to recommend the resort to customers.

5.1.2 K-Means

The application of the k-means method in this study is to cluster two types of data processing, namely the results of processing customer reservation history data with the RFM method and grouping the types of resorts. The first is grouping the k-means method is to create clusters from the results of processing customer reservation history with the RFM method to find the characteristics of the group so that to know the preferences of the customer's and is used as an additional alternative to provide recommendations to customers by finding similarities with other customers. Customer preferences are used to provide resort services recommendations to customers in the future based on characteristics of the resorts he once ordered. Characteristics resort is obtained based on clustering the type of resort

Customer reservation data with rating value has been classified then processed by K-Means method. The number of clusters formed based on high resort prices, moderate resort prices, and low resort prices. The number of clusters (k) for Reservation Data Rancabali Resort formed is 3 ($k = 3$), with the initial centroid point is (1,1) (3,2) (5,1). The initial centroid point for Reservation Data Malabar Resort Clustering is (7,1) (8,9) (9,1). The initial

Commented [A6]: Revision of point number 2

centroid for Reservation Data Ciater & Sukawana Resort is (11,1) (13,2) especially for Ciater and Sukawana created cluster is 2 ($k=2$) because of the type resort small amount. Clustering reservation data using a spreadsheet tool by adding formulas to it. Process of K-Means iteration continues until there is no change in data shift to another cluster. Plotting using a tool from Plotly.

Formula (1) is used to find the value distance (d), where (d_i) is the data to- n , (C_i) is the cluster to- i , and (C_{ia}) is the value of the centroid cluster- i in the data to- i . Variable A is the value of the facility booked customer; variable B is the value of the frequency. The formula (2), (3), (4) example calculated to find the value (d) in first data of customer reservation data history Rancabali that is the customer (RC001) for distance to each cluster.

$$d_i C_i = \sqrt{(A - C_{ia})^2 - (B - C_{ib})^2} \quad (1)$$

$$d_1 C_1 = \sqrt{(1-1)^2 - (1-1)^2} = 0 \quad (2)$$

$$d_1 C_2 = \sqrt{(1-3)^2 - (1-2)^2} = 2.23606797 \quad (3)$$

$$d_1 C_3 = \sqrt{(1-5)^2 - (1-1)^2} = 4 \quad (4)$$

The calculation of formula (2) is a way of determining the distance value (d) in cluster 1 of the 1st data (RC001). In formula (3) is the calculation of distance value (d) in cluster 2 and the formula (4) distance value (d) in cluster 3.

Table 2. The First Iteration Results of Rancabali Customer Cluster

ID Customer	$d_1 C_1$	$d_2 C_2$	$d_3 C_3$	Cluster
RC001	0	2.23606797	4	C1
RC002	0	2.23606797	4	C1
RC003	4	2.23606797	0	C3
RC004	5	3.1622776	1	C3
RC005	0	2.23606797	4	C1
....
RC062	1.41421356	1	3.1622776	C2
RC062	2.23606797	0	2.23606797	C2
RC062	0	2.23606797	4	C1
....

Table 2 is the result of the first iteration of cluster customer by type of resort and rating. Cluster number of each data is determined by comparing the values $d_1 C_1$ to $d_3 C_3$ select the smallest value and make sure the location of the value becomes the value of the cluster.

The K-means method is also used to group types of resorts. This type of resort is a group that is used as a reference to recommend resorts to

customers. The type of resort is recommended based on the position of the resort group the customer has used. For a group of resorts types, the data used only includes the name, number of rooms, the capacity of the room, maximum capacity of the room, price on weekends and prices for facilities and the experiment.

Table 3. Example of Resort Data

Resort Name	Code of Resort Name	Number of Room	Min. Capacity	Max. Capacity	Weekday Price	Weekend Price	Code of Facilities
VIP	R1	1	2	4	500	750	1
Ruang Standar	R2	1	2	4	400	650	1
Bungalow 1	R3	5	10	20	1250	1400	2
Bungalow 2	R4	7	14	28	1550	1850	2
....
Rumah Tahan Gempa	R25	2	4	8	350	350	7

Table 4. Data Group of Resort Facilities

Code	Facilities
1	TV, Hot Water Bathroom, Swimming Pool, Entry Ticket, Breakfast, Living Room
2	TV, Hot Water Bathroom, Swimming Pool, Entry Ticket, Breakfast, Living Room, Kitchen
3	TV, Hot Water Bathroom, Swimming Pool, Breakfast, Living Room, Kitchen
4	TV, Hot Water Bathroom, Breakfast, Living Room, Kitchen
5	TV, Breakfast, Hot Water Bathroom, AC
6	TV, Hot Water Bathroom
7	TV, Living Room, Kitchen, Hot Water Bathroom

Table 3 is that data resort information is converted into numbers to facilitate grouping. Minimum and Maximum Capacity is the amount of capacity for per-person; for example, four that is four-person capacity. Weekend and Weekday Price variables contain values in the rupiah index (IDR) by eliminating the units of thousands of numbers to facilitate the process of clustering. For example, weekday price 350, which means IDR 350,000. Facility data are classified into seven types of facilities because one kind of resort with the other has the same facilities. Table 4 is the list data group of resort facilities.

In data processing with *RapidMiner* need setting parameters to set the number of clusters formed (k). The number of clusters is formed based on high prices, medium prices, and low prices. The grouping resort types grouped into three groups then the parameter k fill with value 3. The max

Commented [A7]: Revision of point number 4

runs parameter represents the maximum number of K-Means is 10. Type of measurement used is NumericalMeasures and numerical sizes using EuclideanDistance calculations with maximum iterations as much as by default is 100 times iterated.

5.2 Result

5.2.1 RFM

The frequency value is obtained based on reservation history data stored in the book from a set of ordering data. Customer reservation history data used reached 100 customer data with various types of resorts used. The resort recommendation scheme gives to customers based on the highest rating.

Table 5. Rancabali Customer Rating

ID Customer	Resort Booked (A)	Frequency/Rating (B)
RC001	1	1
RC002	1	1
RC003	5	1
RC004	6	1
RC005	1	1
RC062	2	2
RC062	3	2
RC062	1	1

Table 5 Rancabali Customer Rating. The data include customer name data, the type of resort booked, and frequency and use as a rating. Column resort booked, "1" is resort type Kidang Kencana, "2" is Ciung Wanara, "3" is Dungus Wangi, "5" is Cottage and "6" is Villa Patenggang.

5.2.2 K-Means

The cluster of each customer based on the results of processing data with K-Means can be seen based on the last iteration. Table 6 is the result of the last iteration of Rancabali customers. The cluster that formed is used to find the characteristics of the group and each customer.

Table 6. The Final Iteration Results of Rancabali Customer Cluster

ID Customer	$d_1 C_1$	$d_2 C_2$	$d_3 C_3$	Cluster
RC001	0	1.362233506	3.484375	C1
RC002	0	1.362233506	3.484375	C1
RC003	4	2.639484822	1.125108502	C3
RC004	5	3.63932413	1.815797109	C3
RC005	0	1.362233506	3.625033674	C1
RC062	1.414213562	1.011118254	3.18937598	C2
RC062	2.236067977	1.140245642	2.490656367	C2
RC062	0	1.362233506	3.625033674	C1

Commented [A8]: Revision of point number 2

Table 6 is the example result of the final iteration of the Rancabali customer cluster. Customers (RC062) enter into two clusters, namely C1 and C2, this is a good thing to make other clusters as an additional alternative to recommending resorts. Figure 1 is a plot of clustering results on customer reservation data of Rancabali resorts. Cluster 1 only contains a group of customers who use the Kidang Kencana resort (1) with a rating value of 1. Cluster 2 is a group of customers who use Ciung Wanara (2) and Dungus Wangi (3) resorts with a rating of 1 and 2. Cluster 3 contains a group of customers who use the Rumah Kayu (4), Cottage (5) and Villa Patenggang (6) with each rating value 1.

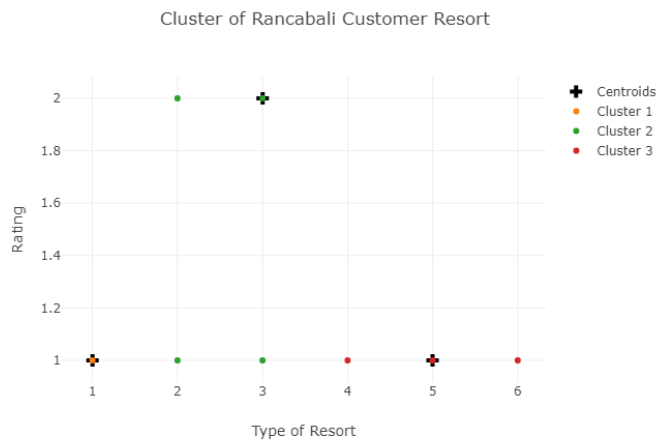


Figure 1. Plot Cluster of Rancabali Customer Resort

Table 7. The final Iteration Result of Malabar Customer Cluster

ID Customer	Facilities Booked (A)	Rating (B)	d_1C_1	d_2C_2	d_3C_3	Cluster
MR001	7	1	0.486504	8.01561	2.217699	C1
MR001	8	1	0.560008	8.01561	1.218916	C1
MR002	9	1	1.546135	8.13941	0.230919	C3
MR003	10	1	2.543119	8.381527	0.787966	C3
MR004	9	2	1.755802	7.158911	0.944013	C3
MR005	9	1	1.546135	8.13941	0.230919	C3

Table 7 is the example result of the final iteration of the Malabar customer cluster. Figure 2 is a plot of the Malabar Customer Cluster. Clusters 1 show three dots representing the value of each data, this is because the data have similarities between one data with another so that the variation of the data there are only three differences. The data in cluster 1 has data groups (7,1) (8,1) and (8,3) and on cluster 2 show two dots consisting of data

(7, 9) and (8,9). While in cluster 3 show five points with data (7,2), (9,2), (9,1) (10,2) and (10,1).

Cluster 1 contains a group of customers using the Rumah Standar Atas (7) with a rating of 1 and a Rumah Standar Bawah (8) with a rating of 1 and 2. Cluster 2, a group of customers with a Rumah Standar Atas (7) with a rating of 9 and Rumah Standat Bawah(8) with a rating of 9. Cluster 3, customers with the Rumah Standar Atas (7) resort group with rating 2, Rumah Kayu (9) with a rating of 2 and 1, Wisma Melati (10) with a rating of 1 and 2.

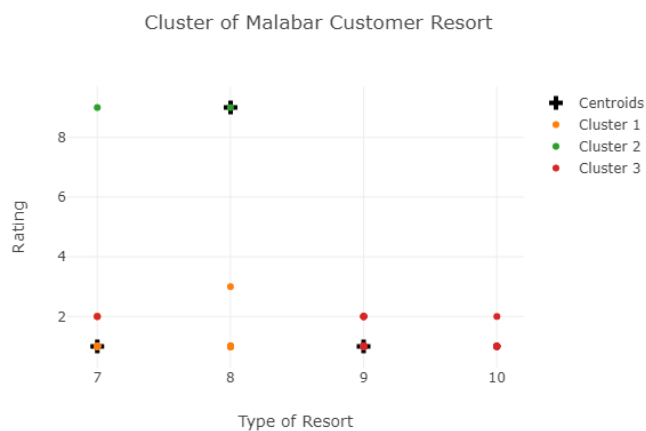


Figure 2. Plot Cluster of Malabar Customer Resort

Table 8. The Final Iteration Result of Sukawana & Ciater Customer Cluster

ID Customer	Facilities Booked (A)	Rating (B)	d_1C_1	d_2C_2	Cluster
SC001	13	1	12.08011	11.98612	C2
SC002	12	1	12.0063	10.98654	C2
SC002	13	1	12.08011	11.98612	C2
SC003	13	1	12.08011	11.98612	C2
SC004	11	1	12.01555	9.987046	C2
SC005	13	1	12.08011	11.98612	C2
SC023	13	5	8.119668	12.52998	C1

In the clustering resort, Sukawana & Ciater grouping clashing is only divided into 2 clusters because the type of resort there are only three types used as the value of the object calculation. Table 8 is the example result of the final iteration of Sukawana & Ciater customer cluster. The clustering results are formed in cluster 1 (C1), the data consists of 11 resorts (Rumah Kayu Atas), 12 (Rumah Kayu Bawah) and 13 (Villa Merah) with a frequency less than six times. While in cluster 2 (C2) only given by data with resort type worth 13 (Villa Merah) with a rating more than six.

Figure 3 is a plot of clustering results on customer reservation data of Sukawana & Ciater resorts. Cluster 1, contains customer group with Villa Merah(11) with a rating of 1 and 2, and Rumah Kayu Atas(12) with rating 1. Cluster 2, contains customer group Rumah Kayu Atas(12) with rating 2, dan 3, and Rumah Kayu Bawah (13) with rating 1, 2, 3, 6, 7 and 15.

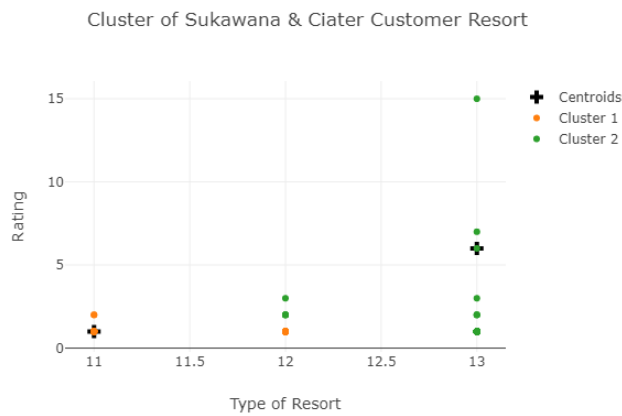


Figure 3. Plot Cluster of Sukawana & Ciater Customer Resort

The result of clustering of resort-type is cluster 0 is 9 data; cluster 1 is 9 data, and cluster 2 is 7 data. Figure 4 results are clustering the group data of the resort type. The data used only include resort name, number of room, a capacity minimum of room, a capacity maximum of room, price on a weekday, and price on weekend and facilities.

Row No.	Id	NAMA RESO...	cluster	Jumlah Kam...	Kapasitas M...	Kapasitas M...	Harga Week...	Harga Week...	Fasilitas
1	1	R1	cluster_0	1	2	4	500	750	1
2	2	R2	cluster_1	1	2	4	400	650	1
3	3	R3	cluster_2	5	10	20	1250	1400	2
4	4	R4	cluster_2	7	14	28	1550	1850	2
5	5	R5	cluster_0	2	4	8	550	750	1
6	6	R6	cluster_0	2	4	8	750	1000	2
7	7	R7	cluster_2	2	4	8	900	1300	2
8	8	R8	cluster_2	2	20	25	1300	1600	3
9	9	R9	cluster_2	3	9	15	1100	1250	4
10	10	R10	cluster_0	2	4	8	800	1150	4
11	11	R11	cluster_0	2	4	8	550	750	4
12	12	R12	cluster_0	2	4	8	800	1150	4
13	13	R13	cluster_0	3	15	25	750	1100	4
14	14	R14	cluster_1	1	2	2	325	475	5
15	15	R15	cluster_1	1	2	2	250	375	5
16	16	R16	cluster_0	2	4	8	700	1000	4
17	17	R17	cluster_2	4	20	50	1200	1600	4
18	18	R18	cluster_2	3	8	50	1000	1600	4
19	19	R19	cluster_0	2	4	8	550	800	4
20	20	R20	cluster_1	2	4	8	400	500	4
21	21	R21	cluster_1	1	2	2	350	350	6
22	22	R22	cluster_1	1	2	2	300	300	6
23	23	R23	cluster_1	1	2	2	250	250	6
24	24	R24	cluster_1	1	2	2	200	200	6
25	25	R25	cluster_1	2	4	8	350	350	7

Figure 4. The Result of Cluster Type Resort

Figure 4 is the group data of the resort type. The most noticeable characteristic of each cluster is that the difference lies in the price. In cluster_0 (cluster 1) consists of resort-type data with prices on a weekday starting at more than IDR 500,000 and less than IDR. 1,000,000. Cluster_1 (cluster 2) consists of a resort-type with a price on a weekday starting from less than IDR 500,000 while on cluster_2 (cluster 3) consists of resort types with prices on a weekday starting from less than IDR 1,000,000 to more than IDR 1,000,000. Figure 5 is a plotting of cluster data of resorts type showing the difference between cluster 0, cluster 1, and cluster 2. The apparent difference is in the price of each resort type.

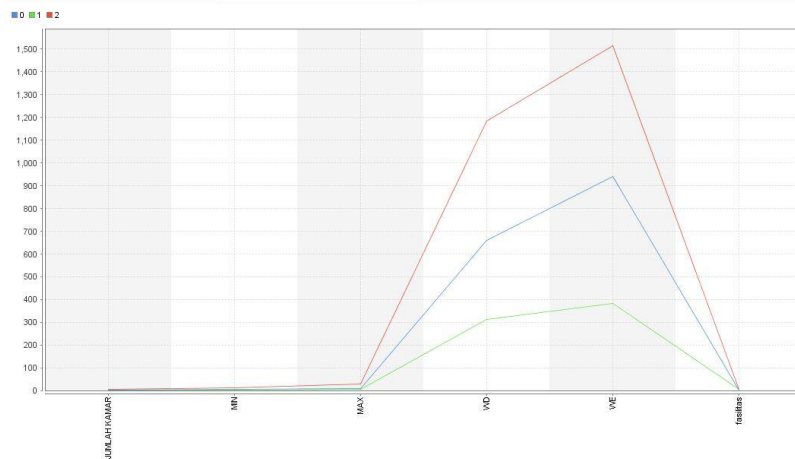
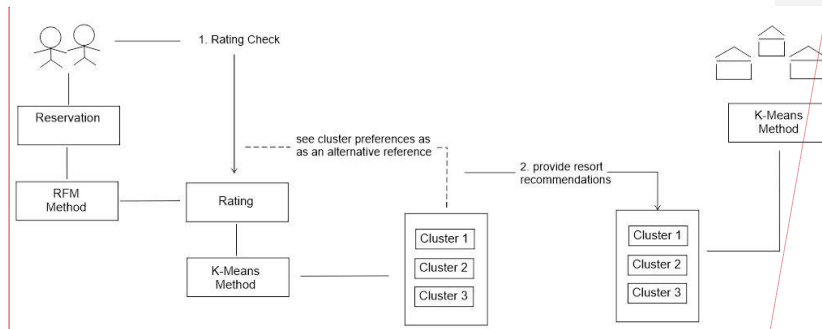


Figure 5. Plotting of Type Resort Clustering

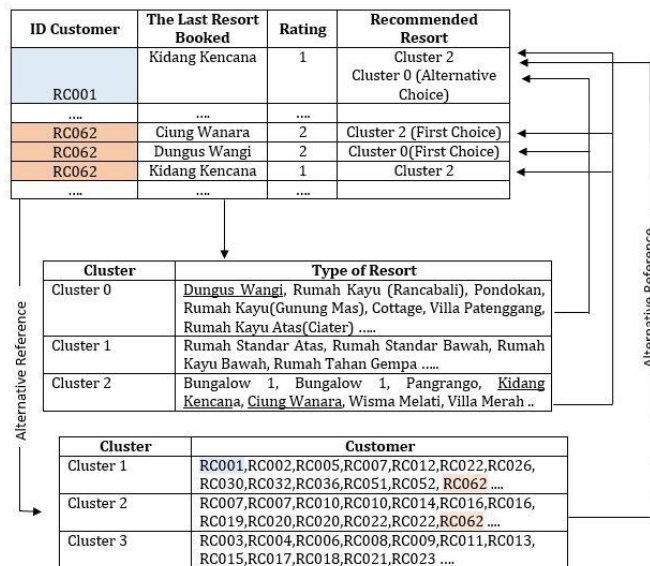
The history customer subscribes data from the data of the customer, and the type of resort is required. Recommendation resort to the customer based on the dominant use in one of the resort types. The recommended type of resort based on the similarity between the types of resorts used with other types of resorts.



Commented [A9]: Revision of point number 3

Figure 6. The flow of Recommendation Schemes

Figure 6 is the flow of the recommendation scheme. Customers get ratings from customer reservation history data, rating as a determinant of the accuracy of recommendations. Customers are divided into three groups to find preferences for resort types; this preference group is used as an additional component of recommendations to find the same criteria as other customers. Types of resorts are grouped to determine resort recommendations to customers precisely according to customer preference criteria.



Commented [A10]: Revision of point number 3

Figure 7. Table of Recommendation Flow

Figure 7 is the table of recommendation flow; for example, Customer (RC001) has a record of resort use of Kidang Kencana (1). Resort Kidang Kencana is part of cluster 2 so that the right recommendation for the Customer (RC001) is a resort that is in cluster 2. As an alternative customer (RC001), it is in a group with the customer (RC062), which is also part of the cluster 2 customer group. So that the customer (RC001) can use cluster 0 on the type of resort. Customer (RC062) has a lot of reservation history, so the priority is the highest rating value among others. Then the members of the group will be recommended to the customer for other examples.

6. CONCLUSION

The scheme made in this study that is the first scheme is customer ever make a reservation then there is reservation history data. Secondly, reservation history data then analyzed into RFM form to know the value of variable frequency. Thirdly, from reservation history then preference classification process here aims to know customer interest to resort-type by using the K-Means method. Next will form a group of customers. Previously resort types have grouped by using the method of K-means also this purpose as a reference to recommend resorts to customers. The fourth scheme of recommendation gives to the customers is to determine the type of resort most rented by the customer then find the type of resort is in the group which is on a cluster of clusters that have formed and members in the group make the resort recommendations to the customer.

A summary of this paper is through a resort-specific design for customers that allow the management to provide the right resort advice to its customers. The company can understand the customer's appetite from reservation history. It is easy to provide information, and for customers, it will be fun for them, this provides value to customer satisfaction that benefits companies for improving customer response. Each customer includes in some different clusters based on reservation history. From the clustering with K-means and simple data, it can be used to design resort recommendations that provide benefits to actors in service and provide satisfaction for the customers.

REFERENCES

- [1] Hongbing Wang, Shizhi Shao, Xuan Zhou, Cheng Wan, Athman Bouguettaya, **Preference Recommendation for Personalized Search**, *Knowledge-Based Systems*, 2016.
- [2] Kai Zhang, Keqiang Wang, Xiaoling Wang, Cheqing Jin, Aoying Zhou, **Hotel Recommendation based on User Preference Analysis**, pp. 134-138, 2015.
- [3] Le Hoang Son, **HU-FCF++: A Novel Hybrid Method for The New User Cold-Start Problem in Recommender Systems**, *Engineering Applications of Artificial Intelligence*, Vol. 41, pp. 207-222, 2015.

- [4] Syafrial Fachri Pane, Rolly Maulana Awangga, Bayu Rahmad Azhari, **Qualitative Evaluation of RFID Implementation Warehouse Management System**, *TELKOMNIKA (Telecommunication Comput. Electron. Control*, Vol. 16, 2018.
- [5] Rolly Maulana Awangga, Syafrial Fachri Pane, Khaera Tunnisa, Iping Supriana Suwardi, **K Means Clustering and Meanshift Analysis for Grouping the Data of Coal Term in Puslitbang tekMIRA**, *TELKOMNIKA (Telecommunication Comput. Electron. Control*, Vol. 16, 2018.
- [6] Keng-Pei Lin, Chia-Yu Lai, Po-Cheng Chen, San-Yih Hwang, **Personalized Hotel Recommendation Using Text Mining and Mobile Browsing Tracking**, *IEEE International Conference on System, Man, and Cybernetics*, pp. 191-196, 2016.
- [7] Kanae Matsui, Hanjong Choi, **A Recommendation System with Secondary usage of HEMS Data for Products based on IoT Technology**, *International Symposium on Networks, Computers and Communications, ISNCC*, 2017.
- [8] Imran Memon, Ling Chen, Abdul Majid, Mingqi Lv, Ibrar Hussain, Gencai Chen, **Travel Recommendation Using Geo-tagged Photos in Social Media for Tourist**, *Wireless Personal Communications*, Vol. 80, No.4, pp. 1347-1362, 2015.
- [9] Idir Benouaret, Dominique Lenne, **A Composite Recommendation System for Planning Tourist Visits**, *International Conference on Web Intelligence*, pp. 626-631, 2016.
- [10] Blake Hallinan, Ted Striphas, **Recommended for you : The Netflix Prize and The Production of Algorithmic Culture**, *New Media & Society*, Vol. 18, No. 1, pp.117-137, 2016.
- [11] Jenet Manyi Agbor, **The Relationship between Customer Satisfaction and Service Quality: a study of three Service sectors in Umea**, UMEA Universiti, 2011.
- [12] William Wei Song, Chenlu Lin, Anders Forsman, Anders Avdic, Leif Akerblom, **An Euclidean Similarity Measurement Approach for Hotel Rating Data Analysis**, *IEEE 2nd International Conference on Cloud Computing and Big Data Analysis*, pp. 293-298, 2017.
- [13] Koji Takuma, Junya Yamamoto, Sayaka Kamei, Satoshi Fujita, **A hotel Recommendation System based on Reviews: What Do You Attach Importance To?**, *Fourth International Symposium on Computing and Networking*, pp. 710-712, 2017.
- [14] Huiming Wang, Nianlong Luo, **Collaborative Filtering Enhanced by User Free-Text Reviews Topic**, *International Conference on Information and Communications Technologies (ICT)*, 2014
- [15] Anbazhagan Mahadevan, Michael Arock, **Credible User-Review Incorporated Collaborative Filtering for Video Recommendation System**, *International Conference on Intelligent Sustainable Systems (ICISS)*, pp. 375-379, 2017.

- [16] Shahriar Badsha, Xun Yi, Ibrahim Khalil, **A Practical Privacy-Preserving Recommender System**, *Data Science and Engineering*, Vol. 1, pp. 161–177, 2016.
- [17] Zhou Zhao, Deng Cai, Xiaofei He, Yueting Zhuang, **User Preference Learning for Online Social Recommendation**, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 28, No. 9, pp. 2522–2534, 2016.
- [18] Zhou Zhao, Qifan Yang, Hanqing Lu, Tim Weninger, Deng Cai, Xiaofei He, Yueting Zhuang, **Social-Aware Movie Recommendation via Multimodal Network Learning**, *IEEE Transactions on Multimedia*, Vol. 20, No. 2, pp. 430–440, 2018.
- [19] Ihsan Topalli, Selcuk Kilinc, **Modelling User Habits and Providing Recommendations based on The Hybrid Broadcast Broadband Television using Neural Networks**, *IEEE Transactions on Consumer Electronics*, Vol. 62, no. 2, pp. 182–190, 2016.
- [20] K. Kesorn, W. Juraphanthong, A. Salaiwarakul, **Personalized Attraction Recommendation System for Tourists Through Check-In Data**, *IEEE Access*, vol. 5, pp. 26703–26721, 2017.
- [21] Xueming Qian, He Feng, Guoshuai Zhao, Tao Mei, **Personalized Recommendation Combining User Interest and Social Circle**, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 26, No.7, pp. 1763–1777, 2014.
- [22] Zhiyang Jia, Wei Gao, Yuting Yang, Xu Chen, **User-based Collaborative Filtering for Tourist Attraction Recommendations**, *IEEE International Conference on Computational Intelligence & Communication Technology (CICIT)*, pp. 22–25, 2015.
- [23] Junge Shen, Cheng Deng, Xinbo Gao, **Neurocomputing Attraction recommendation: Towards personalized tourism via collective intelligence**, *Neurocomputing*, Vol. 173, pp. 789–798, 2016.
- [24] Michalis Korakakis, Phivos Mylonas, Evaggelos Spyrou, **Xenia: A Context Aware Tour Recommendation System Based on Social Network Metadata Information**, *11th International Workshop on Semantic and Social Media Adaption and Personalization (SMAP)*, pp. 59–64, 2016.
- [25] Chin-I Lee, Tse-Chih Hsia, Hsiang-Chih Hsu, Jing-Ya Lin, **Ontology-Based Tourism Recommendation System**, *4th International Conference on Industrial Engineering and Applications*, pp. 376–379, 2017.
- [26] Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, Shanika Karunasekera, **Personalized Tour Recommendation Based on User Interests and Points of Interest Visit Durations**, *International Joint Conference on Artificial Intelligence (IJCAI)*, Vol. 15, pp. 1778–1784, 2015.
- [27] Chenyi Zhang, Ke Wang, **POI Recommendation Through Cross-Region Collaborative Filtering**, *Knowledge and Information Systems*, Vol. 46, No. 2, pp. 369–387, 2017.
- [28] Rachid Aid Daoud, Belaid Bouikhalene, Abdellah Amine, Rachid LBIBB,

- Combining RFM Model and Clustering Techniques for Customer Value Analysis of A Company Selling Online, *IEEE 12th International Conference of Computer Systems and Applications (AICCSA)*, pp. 1-6, 2015.
- [29] Jo Ting Wei, Shih-Yen Lin, You-Zhen Yang, Hsin-Hung Wu, **Applying Data Mining and RFM Model to Analyze Customers' Values of A Veterinary Hospital**, *International Symposium on Computer, Consumer and Control*, pp. 481-484, 2016.
- [30] Kristof Coussement, Filip A.M. Van den Bossche, Koen W. De Bock, **Data Accuracy's Impact on Segmentation Performance: Benchmarking RFM Analysis, Logistic Regression, and Decision Trees**, *Journal of Business Research*, Vol. 67, No. 1, pp. 2751-2758, 2014.
- [31] Ferdi Yusuf, **Pembangunan Sistem Informasi Customer Relationship Management di Koperasi Pegawai dan Pensiunan PT. Pos Indonesia (KOPPOS)**, *Jurnal Ilmiah Komputer dan Informatika (KOMPUTA)*, Vol. 1, No. 1, 2016.
- [32] Michele Amoretti, Laura Belli, Francesco Zanichelli, **UTravel : Smart Mobility with A Novel User Profiling and Recommendation Approach**, *Pervasive and Mobile Computing*, Vol. 38, pp. 474-489, 2017.
- [33] Merlinda Sumardi, Jufery, Frenky, Rini Wongso, Ferdinand Ariandy Luwinda, **"TripBuddy" Travel Planner with Recommendation based on User ' s Browsing Behaviour**, *International Conference on Computer Science and Computational Intelligence (ICCCSI)*, Vol. 116, pp. 326-333, 2017
- [34] Shaoqing Wang, Cuiping Li, Kankan Zhao, Hong Chen, **Context-Aware Recommendations with Random Partition Factorization Machines**, *Data Science and Engineering*, Vol. 2, No. 2, pp. 125-135, 2017.
- [35] Zan Wang, Xue Yu, Nan Feng, Zhenhua Wang, **An Improved Collaborative Movie Recommendation System using Computational Intelligence**, *Journal of Visual Languages and Computing*, Vol. 25, pp. 667-675, 2014.