SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171

Consumer Segmentation of Emina Cosmetics Optimal and Relevant Approach of RFM+Lifetime Analysis

Nabil Rakha Dwitya[1], Wirawan Istiono[2]

Universitas Multimedia Nusantara, Tangerang, Indonesia

Author Email: nabil.dwitya@student.umn.ac.id[1], wirawan.istiono@ umn.ac.id[2]

Abstract. Buyers are the most crucial entities for companies selling products, including PT Paragon Technology and Innovation. PT Paragon Technology and Innovation is a cosmetics company that oversees well-known brands such as Wardah, Emina, MakeOver, and Kahf. It is essential for this company to understand the characteristics of its buyers who purchase their products, and one way to achieve this is by conducting consumer segmentation. This consumer segmentation is carried out on customers who have purchased Emina products from March 2021 to March 2023, using three types of RFM analysis approaches: vanilla RFM analysis, RFM+Lifetime, and RFM/Lifetime, which are then grouped using the K-Means algorithm. Through the implementation of this consumer segmentation, the company can gain a deeper understanding of its buyers' behavior towards the products they offer, thereby enhancing business processes and marketing efforts. The consumer segmentation has been completed with the finding that out of the three types of RFM analysis approaches employed for consumer segmentation, the RFM+Lifetime approach is the most effective and relevant one, resulting in four categories: Make Up, Face Care, Others, and General. The Make Up category further consists of five segments, while each of the other categories contains four segments.

Keywords: Consumer Segmentation, K-Means Algorithm, RFM Analysis, RFML Analysis.

1. Introduction

Certain organizations utilize Customer Relationship Management (CRM) systems in order to discern and recognize consumers who have significant value [1], [2]. client Relationship Management (CRM) is a strategic approach adopted by businesses to cultivate client loyalty and establish enduring, profitable partnerships [3]. Moreover, implementing CRM increases customer value and business profits. One crucial approach to CRM is consumer segmentation, which involves grouping consumers based on similar characteristics [4]. This consumer segmentation technique enables businesses to identify the most profitable customers [5], [6]. By developing and providing desired products and services, businesses can foster customer loyalty, satisfaction among existing customers, and attract new consumers [7].

Often, age, gender, product preferences, and expenditure patterns are examples of characteristics used to group consumer data into distinct segments, known as consumer segmentation. This consumer segmentation method is effective in determining a company's market share. Subsequently, companies can utilize consumer data to plan improvements, run attractive promotions, or develop new products. By grouping consumers, companies gain a better understanding and identification of their target audience. In other words, besides grouping consumers, the goal of segmentation is to tailor products, services, and marketing messages for each segment/group [4], [8].

In this case, the process of consumer segmentation is applied based on three types of RFM analysis approaches: vanilla RFM, RFM+Lifetime, and RFM/Lifetime. RFM analysis itself examines three variables: Recency, Frequency, and Monetary [9], [10]. Recency represents the last time a consumer made a purchase, Frequency is the number of transactions made by the consumer, and Monetary refers to the amount spent by the consumer [11], [12]. RFM+Lifetime and RFM/Lifetime are additional approaches that initially rely on the three RFM variables but add one more variable, Lifetime. Lifetime indicates the duration since the consumer's first purchase [13]. From these three or four factors, specific characteristics are derived and labeled. Typically, companies manually label consumers based on observable shopping habits [14], which may introduce errors or biases when done by humans.

SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171

Consequently, the function of computers in consumer grouping becomes crucial. Using unsupervised learning, computers group data based on a variety of variables [15], [16]. Unsupervised learning aids in modeling consumer behavior and personalities that may have been previously overlooked [17]. One method of unsupervised learning-based grouping is using the K-Means algorithm, a commonly used clustering algorithm [18]. The K-Means algorithm is frequently used in consumer segmentation to organize data into meaningful clusters for visualization purposes. K-Means is a data mining technique that clusters data by evaluating their similarity [15], [19].

This consumer segmentation innovation is what PT Paragon Technology and Innovation seeks to implement. The company has not yet employed such consumer segmentation. Therefore, this consumer segmentation is necessary to support the business and marketing needs of the company. The segmentation is performed on consumers who purchased Emina products from 2021 to March 2023. The consumer data is obtained from purchase transactions on the Shopee application.

2. Methodology

In conducting consumer segmentation for buyers of Emina products for PT Paragon Technology and Innovation, there are three different approaches to RFM analysis, namely vanilla RFM, RFM + Lifetime, and RFM/Lifetime. All of these approaches follow the same process or flow, as depicted in Figure 1.

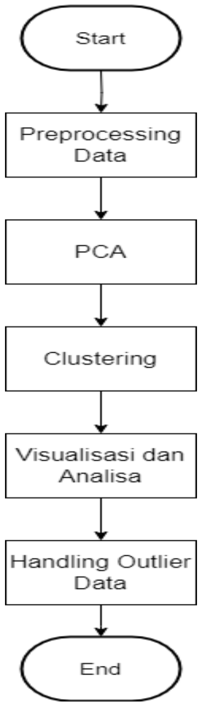


Figure 1. Consumer Segmentation Flow

Figure 1 illustrates the process flow, starting with data preprocessing, followed by principal component analysis (PCA), then moving on to the clustering stage, visualizing and analyzing the clustering results, and finally handling outlier data

2.1 Preprocessing Data

The data preparation process for conducting consumer segmentation encompasses three main aspects: data selection, data cleansing, and data transformation.

2.1.1 Data Cleaning pertama

Data cleaning is necessary to reduce unnecessary data rows. In this case, data reduction is performed on the second dataframe, namely "emina," for rows that have Odoo Code values that do not exist in the first dataframe, "cat." Data cleaning is achieved by removing the Odoo Code from the "emina" dataframe, as shown in Figure 2.

SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171

```
list = ['P4001','P2290','P1961','P2317','P1987','P2639','P1978','P2164','P1988','P2420','P3096','P3095','P3068','P3094',
        'P3170','P3264','P3265','P4238','P4264','P4329','P4412','P4599','P4602','P4528','P4543','P3113','P4426','P3112',
        'P2223','P4060','P3160','P4215','P1848','P1964','P1890','P2156','P1846','P1960','P1736','P3017','T454','P1859',
        'P2594','P4413','04045']
for i in list:
    emina = emina[emina['Odoo Code'] != i]
```

Figure 2. Delete Odoo Code on Emina Dataframe

Figure 2 indicates that many Odoo Code values in the "emina" dataframe were removed. Besides the reason that these Odoo Code values do not exist in the "cat" dataframe, another objective is to enable the merging or combination of both dataframes for further processing.

2.1.2 Data Selection

Data selection is performed by merging the two dataframes based on the Odoo Code values, resulting in a new dataframe that contains all the columns present in both the "cat" and "emina" dataframes. The outcome of merging the two dataframes can be observed in Figure 3.

```
merged_emina = emina.merge(cat,on=["Odoo Code"], how='inner')
merged_emina.head()
```

	Consumer ID	Invoice Number	Order Time	Order Status	Odoo Code	Sales	category
0	c2gtaGFyc2FmaXRyaWk=	MjMwMjI3M1VWNTdIRk4=	2023-02-27 07:50:15.000 +0700	Completed	3537	37500	MAKE UP
1	c2gtaGFyc2FmaXRyaWk=	MjMwMjI3M1VWNTdIRk4=	2023-02-27 07:50:15.000 +0700	Completed	3537	37500	MAKE UP
2	c2gtZGVzaWhvbWJpbmcy	MjIwNDA5NTc2QTAwN0o=	2022-04-09 18:58:46.000 +0700	Completed	3537	37500	MAKE UP
3	c2gtYXFpbGFhbGZpNzk3	MjIwNjI1UzkwMVZTSjQ=	2022-06-26 05:04:20.000 +0700	Completed	3537	37500	MAKE UP
4	c2gtbm92ZWxzaGlmZmEyMg==	MjMwMzAzRlFXVktCSjE=	2023-03-04 01:11:35.000 +0700	Completed	3537	37500	MAKE UP

Figure 3. Merge the Two Dataframes

As seen in Figure 3, there is now a new dataframe called "merged emina," which contains a combination of columns from both the "cat" and "emina" dataframes. The purpose of merging these two dataframes is to determine the transactions made by consumers for purchasing specific product categories. This merged dataframe serves as the continuation of the segmentation process. With the current dataframe, which already includes a "category" column, consumer segmentation can be performed based on the existing categories. However, in this case, the categories to be separated are "Make Up" and "Face Care," while the other categories will be combined into a single category named "Others." "Make Up" and "Face Care" are the most frequently purchased products, making them the main categories in this segmentation process to obtain data specifically for "Make Up," "Face Care," and "Others" categories.

MakeUp Segmentation

```
eminaMakeUp = merged_emina[merged_emina['category'] != 'BODY CLEANSING']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'BODY MOISTURIZER']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'FACE CARE']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'FRAGRANCE']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'PACKAGE BEAUTY']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'OTHERS']
eminaMakeUp.head()
```

	Consumer ID	Invoice Number	Order Time	Order Status	Odoo Code	Sales	category
0	c2gtaGFyc2FmaXRyaWk=	MjMwMjI3M1VWNTdIRk4=	2023-02-27 07:50:15.000 +0700	Completed	3537	37500	MAKE UP
1	c2gtaGFyc2FmaXRyaWk=	MjMwMjI3M1VWNTdIRk4=	2023-02-27 07:50:15.000 +0700	Completed	3537	37500	MAKE UP
2	c2gtZGVzaWhvbWJpbmcy	MjIwNDA5NTc2QTAwN0o=	2022-04-09 18:58:46.000 +0700	Completed	3537	37500	MAKE UP
3	c2gtYXFpbGFhbGZpNzk3	MjIwNjI1UzkwMVZTSjQ=	2022-06-26 05:04:20.000 +0700	Completed	3537	37500	MAKE UP
4	c2gtbm92ZWxzaGlmZmEyMg==	MjMwMzAzRlFXVktCSjE=	2023-03-04 01:11:35.000 +0700	Completed	3537	37500	MAKE UP

Figure 4. Get Make-Up Data

SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171

MakeUp Segmentation

```
eminaMakeUp = merged_emina[merged_emina['category'] != 'BODY CLEANSING']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'BODY MOISTURIZER']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'FACE CARE']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'FRAGRANCE']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'PACKAGE BEAUTY']
eminaMakeUp = eminaMakeUp[eminaMakeUp['category'] != 'OTHERS']
eminaMakeUp.head()
```

	Consumer ID	Invoice Number	Order Time	Order Status	Odoo Code	Sales	category
0	c2gtaGFyc2FmaXRyaWk=	MjMwMjI3M1VWNTdIRk4=	2023-02-27 07:50:15.000 +0700	Completed	3537	37500	MAKE UP
1	c2gtaGFyc2FmaXRyaWk=	MjMwMjI3M1VWNTdIRk4=	2023-02-27 07:50:15.000 +0700	Completed	3537	37500	MAKE UP
2	c2gtZGVzaWhvbWJpbmcy	MjIwNDA5NTc2QTAwN0o=	2022-04-09 18:58:46.000 +0700	Completed	3537	37500	MAKE UP
3	c2gtYXFpbGFhbGZpNzk3	MjIwNjI1UzkwMVZTSjQ=	2022-06-26 05:04:20.000 +0700	Completed	3537	37500	MAKE UP
4	c2gtbm92ZWxzaGlmZmEyMg==	MjMwMzAzRlFXVktCSjE=	2023-03-04 01:11:35.000 +0700	Completed	3537	37500	MAKE UP

Figure 5. Get Face-Care Data

Others Segmentation

```
eminaOthers = merged_emina[merged_emina['category'] != 'FACE CARE']
eminaOthers = eminaOthers[eminaOthers['category'] != 'MAKE UP']
eminaOthers.head()
```

	Consumer ID	Invoice Number	Order Time	Order Status	Odoo Code	Sales	category
1709901	c2gtdHJpeWI4Y2hpbGw=	MjIwNDI3TTJGV0dQMzU=	2022-04-27 06:31:17.000 +0700	Completed	4392	113000	PACKAGE BEAUTY
1709902	c2gtYXprYTE5MDMxMA==	MjMwMzExNVRWVkFYVjg=	2023-03-11 16:24:45.000 +0700	Completed	4392	113000	PACKAGE BEAUTY
1709903	c2gtaWthbWxhdGhpZmFo	MjIwNzI4S1gzMTY1QzY=	2022-07-28 19:25:50.000 +0700	Completed	4392	113000	PACKAGE BEAUTY
1709904	c2gtZmFyYWhkeWFoMTQ=	MjMwMzA1TU4yRjJQTVA=	2023-03-06 00:03:46.000 +0700	Completed	4392	113000	PACKAGE BEAUTY
1709905	c2gtcHBpdHJhbmlhYQ==	MjMwMzA0R0ZEU1RKTVQ=	2023-03-04 08:12:05.000 +0700	Completed	4392	113000	PACKAGE BEAUTY

Figure 6. Get Others Data

From Figure 4, it can be observed that to obtain the "Make Up" data, all categories except "Make Up" need to be removed. The same applies to obtaining the "Face Care" data, as shown in Figure 5. And to obtain the "Others" data, only the "Make Up" and "Face Care" data need to be removed, as depicted in Figure 6. In addition to these three segmented categories, there is also a general category known as the "Others" category. The "Others" category does not consider the specific product categories, so the original "emina" dataframe is used for this segmentation. The general consumer segmentation aims to provide an overview of the entire dataset. As a result, there will be four dataframes used for consumer segmentation in each approach: "eminaMakeUp," "eminaFaceCare," "eminaOthers," and "emina." Once the data selection process is complete, the next step is data transformation.

2.1.3 Data Transformation

Data transformation is used to obtain the appropriate features/columns based on the three analysis approaches. For vanilla RFM, data transformation is carried out solely to obtain the "recency," "frequency," and "monetary" columns. This approach does not involve data transformation for the "lifetime" column. In contrast, the RFM+Lifetime approach involves data transformation to obtain the "recency," "frequency," "monetary," and "lifetime" columns. As for the RFM/Lifetime approach, data transformation is similar to the RFM+Lifetime approach, but the "lifetime" column will be used as a divisor for the "monetary" and "frequency" columns, resulting in only three columns.

2.1.4 Second Data Cleaning

After performing data transformation to obtain the required columns for each approach, the next step is to conduct data cleaning. Data cleaning is applied to all dataframes in each approach. The purpose of this data cleaning stage is to eliminate noise or outlier data present in each dataframe for each approach. By removing these outlier data points, the data used for clustering in the subsequent steps becomes more valid and consistent. To identify whether the data used contains outlier data, boxplots are used, as shown in Figure 7.

SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171

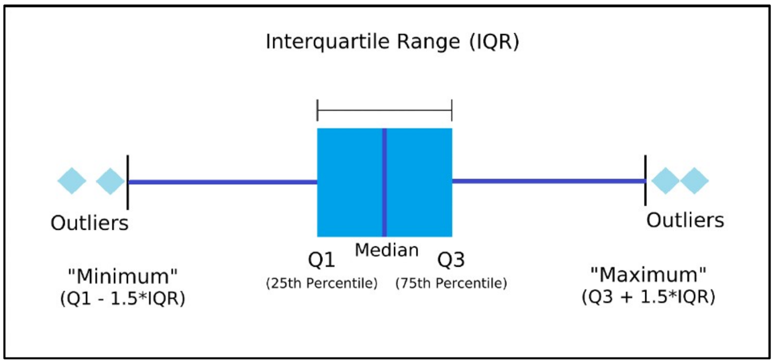


Figure 7. Outlier Data Example

Figure 7 provides an example of using a boxplot to identify outlier data. As observed from Figure 7, outlier data points lie outside the minimum range, which is 1.5 times the interquartile range (IQR) below the first quartile, and the maximum range, which is 1.5 times the IQR above the third quartile. Therefore, data points that fall within the minimum and maximum ranges are considered non-outliers.

2.1.5 Second Data Transformation

After performing data cleaning to remove outlier data from all dataframes in each approach, the next step is data transformation. This time, the data transformation involves standardization. Standardization is performed to ensure that the data in the dataframes have the same range of values, avoiding any dominant or excessive features over others. This is particularly evident in features like "monetary" and "frequency." The "monetary" feature may have values in the tens to hundreds of thousands, while the "frequency" feature may have values in the units or tens. Standardization is carried out using the StandardScaler function, which automatically standardizes the applied dataframe.

2.2 Principal Component Analysis

In conducting this consumer segmentation, there is a consideration regarding whether the segmentation should be separated based on categories such as "Make Up," "Face Care," and "Others," or if it should be directly combined into one general category. This consideration arises because the behavior of individuals in using "Make Up," "Face Care," and "Others" may differ. Therefore, there is a question about whether to perform a separate segmentation for these categories. To address this, Principal Component Analysis (PCA) can be used to examine the correlation between the values of "monetary," "frequency," and "recency" for the "Make Up," "Face Care," and "Others" categories. If these variables are highly correlated, then there may not be a need to divide the consumer segmentation based on these categories. However, if they are not highly correlated, then it would be necessary to separate the consumer segmentation for these categories. Principal Component Analysis is conducted by combining all columns/features within each category. For the vanilla RFM and RFM/Lifetime approaches, there will be nine columns, while for the RFM+Lifetime approach, there will be twelve columns.

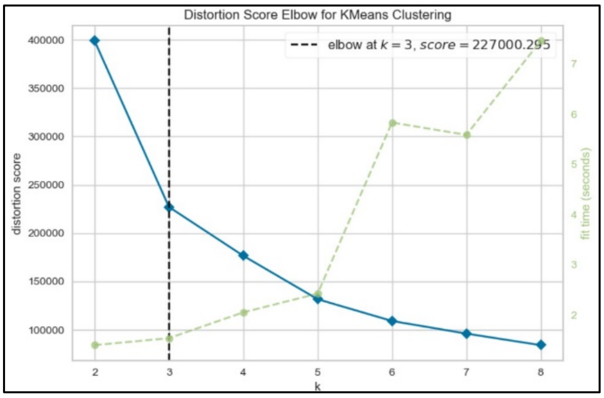
3. Result

Based on the previous Principal Component Analysis, clustering or segmentation will indeed be separated based on the categories mentioned earlier. Therefore, there will be four segmentations for each approach. These four segmentations will be for the categories "Make Up," "Face Care," "Others," and the general category. Each approach will have its respective segments within these categories, allowing for a more detailed and tailored consumer segmentation.

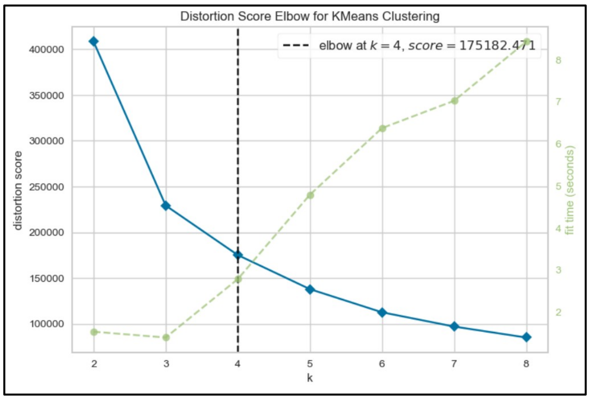
3.1 Consumer Segmentation using the Vanilla RFM Approach

The segmentation process using the K-means algorithm requires initializing the number of clusters or centroids to be used. The method to select the appropriate number of clusters is by using the elbow method. The elbow method is a technique for choosing the right number of clusters by calculating the sum of squared distances (SSD) from each data point to the centroid or center of each cluster. The results of using the elbow method for each category can be seen in Figure 8.

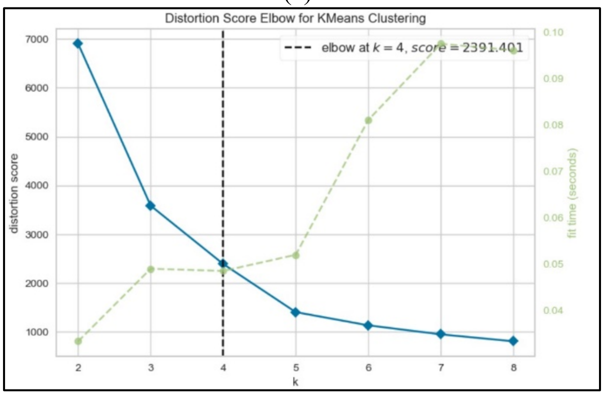
SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171



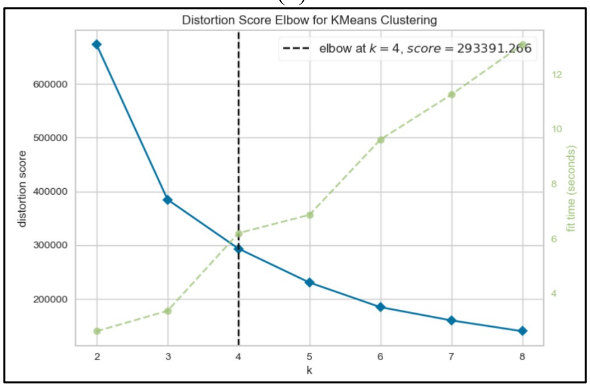
(a)



(b)



(c)



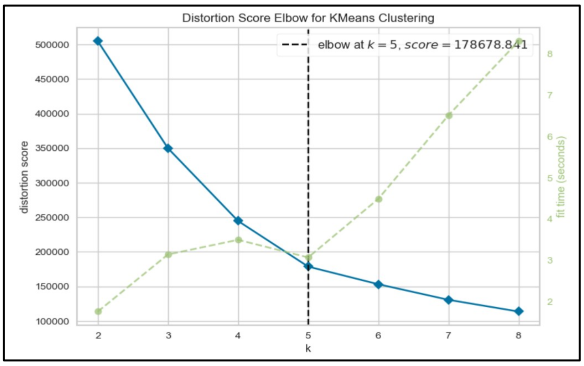
(d)

Figure 8. Elbow Method Result For Each Category With Vanila RFM Approach

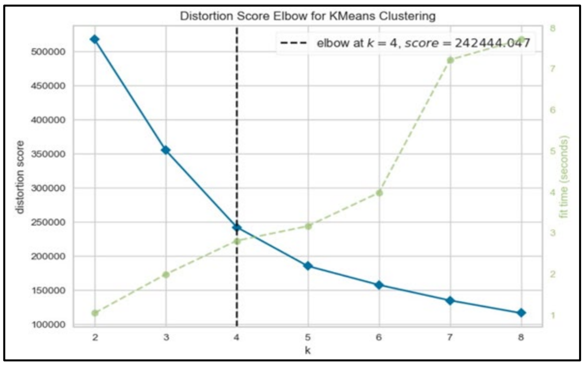
The selection of the desired number of clusters using the elbow method involves identifying a point on a given value of "k" or centroid, where the point appears to form an "elbow" shape. As shown in Figure 8a, the elbow point is at the value of three, indicating that the number of clusters for segmenting the "Make Up" category is three. Similarly, in Figure 8b, the elbow point is at the value of four, meaning that the number of clusters for segmenting the "Face Care" category is four. The same principle applies to the segmentation of the "Others" category, as seen in Figure 8c, where the number of clusters used is four. For the general category in Figure 8d, the elbow point is at the value of four, indicating that the number of clusters to be used will also be four.

3.2 Consumer Segmentation RFM+Lifetime Approach

Similar to the consumer segmentation using the vanilla RFM approach, the consumer segmentation using the RFM+Lifetime approach also requires initializing the appropriate number of clusters using the k-means algorithm by employing the elbow method. The results of using the elbow method for each category can be seen in Figure 9.

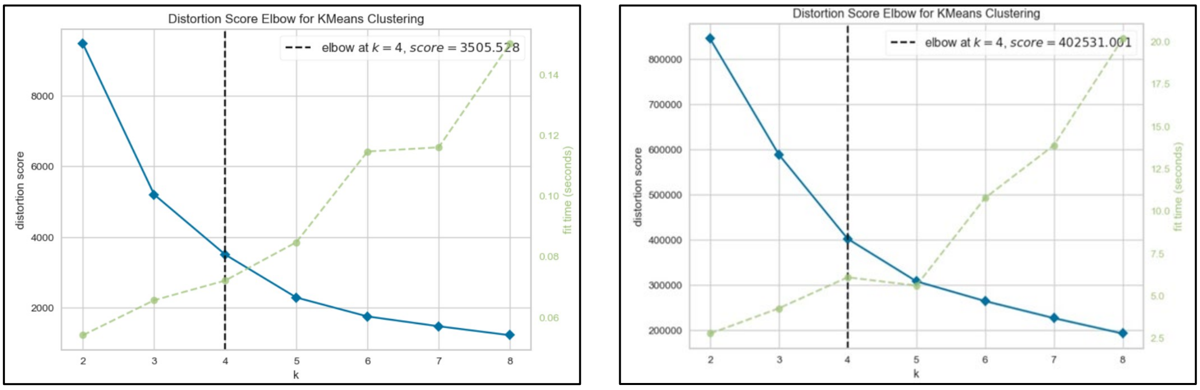


(a)



(b)

SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171



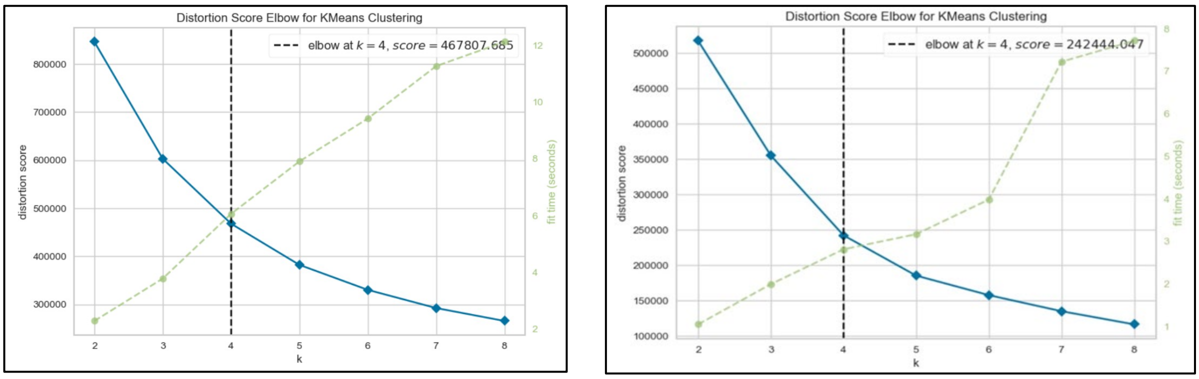
(c) (d)

Figure 9. Elbow method result for each category with RFM+Lifetime Approach

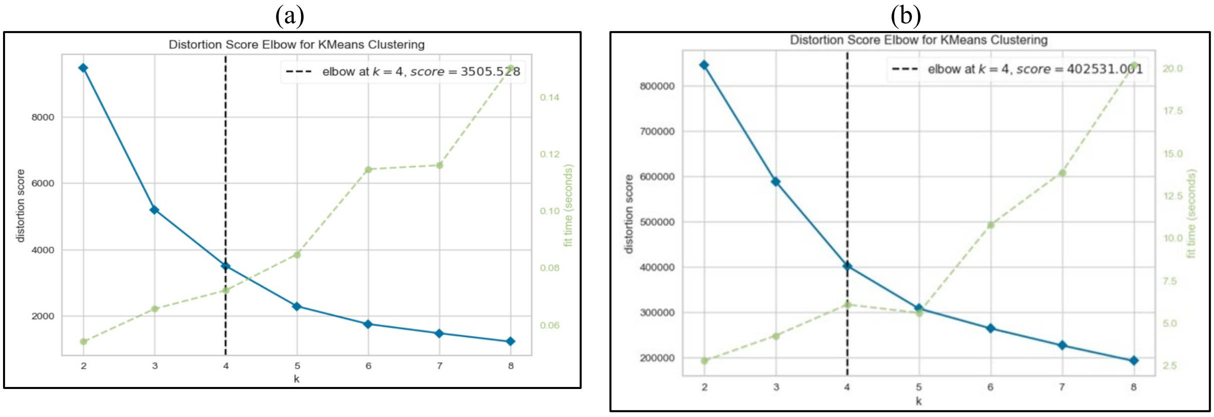
As shown in Figure 9a, the elbow point is at the value of five, indicating that the number of clusters for segmenting the "Make Up" category is five. Similarly, in Figure 9b, the elbow point is at the value of four, meaning that the number of clusters for segmenting the "Face Care" category is four. The same principle applies to the segmentation of the "Others" category, as seen in Figure 9c, where the number of clusters used is four. For the general category in Figure 9d, the elbow point is at the value of four, indicating that the number of clusters to be used will also be four.

3.3 Consumer Segmentation RFM/Lifetime Approach

Just like the two previous approaches, performing segmentation using the K-Means algorithm requires initializing the appropriate number of clusters using the elbow method for each category. The results of using the elbow method for each category can be seen in Figure 10.



(a) (b)



(c) (d)

Figure 10. Elbow method result for each category with RFM/Lifetime Approach

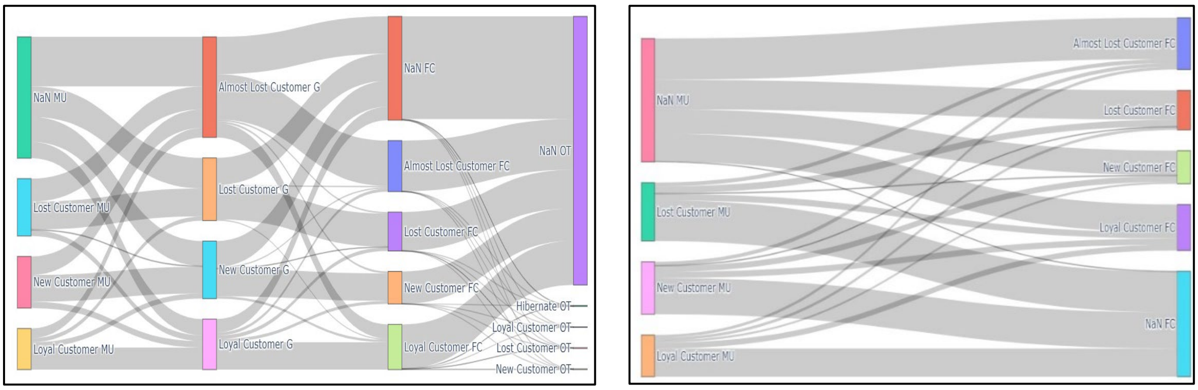
SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171

3.4 Visualization of Segmentation Results

After performing clustering in the previous stage, the next step is to examine the results of the created clusters by visualizing them. The visualizations for each approach are as follows.

A. Visualization of the Vanilla RFM Approach

After observing the visualization for the vanilla RFM approach, it is evident that the frequency value does not play a significant role in labeling the formed segments. This is due to all the visualized data having a frequency value of one. Therefore, the vanilla RFM approach in this case is considered less effective as it relies solely on the monetary and recency values. Additionally, an additional visualization using a Sankey diagram to observe the distribution of cluster labels across all categories can be seen in Figure 11.



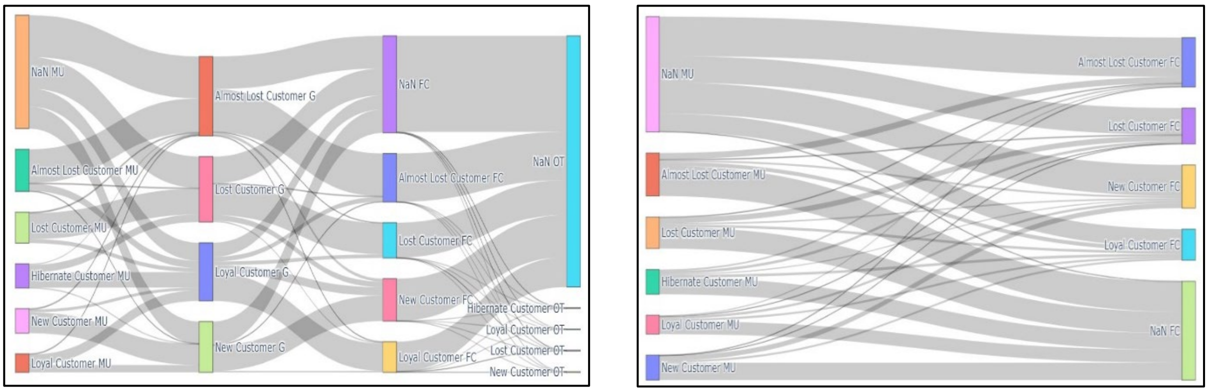
(a) (b)

Figure 11. Sankey Diagram Vanilla RFM 1 and RFM 2

From Figure 11a, it can be observed that consumers who are loyal to the "Make Up" category are mostly also loyal to the general category. Similarly, consumers who are loyal to the "Face Care" category are also mostly loyal to the general category. Additionally, from Figure 11b, it is evident that consumers who are loyal to the "Make Up" category mostly do not purchase products from the "Face Care" category. On the other hand, consumers who are loyal to the "Face Care" category mostly do not purchase products from the "Make Up" category. These insights represent a small portion of the information that can be derived from both of these figures.

B. Visualization of the RFM+Lifetime Approach

The RFM+Lifetime approach proves to be the most effective due to the necessity of additional features such as the lifetime feature, which adds value in labeling the segments since the frequency feature cannot be considered. The utilization of the Sankey diagram, as in the previous approach, can be observed in Figure 12.



(a) (b)

Figure 12. Sankey Diagram RFM+Lifetime1 and RFM+Lifetime2

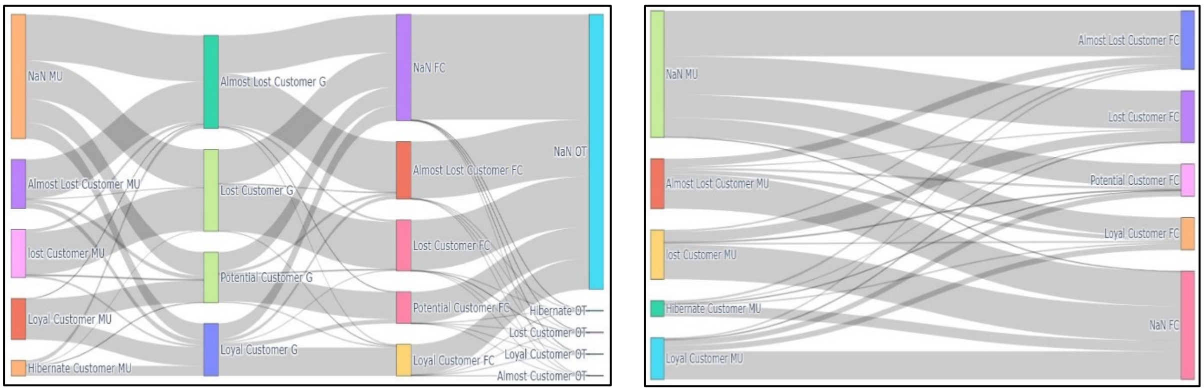
From Figure 12a, it can be observed that consumers who are loyal to the "Make Up" category are mostly also loyal to the general category. Similarly, consumers who are loyal to the "Face Care" category are almost all loyal to the general category as well. Additionally, from Figure 12b, it is evident that consumers who are loyal to the "Make Up" category mostly do not purchase products from the "Face Care" category. On the other hand,

SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171

consumers who are loyal to the "Face Care" category mostly do not purchase products from the "Make Up" category. This information is similar to the findings of the previous approach.

C. Visualization of the RFM/Lifetime Approach

The RFM/Lifetime approach proves to be less appropriate when applied to this consumer segmentation case. This is due to the bias that occurs in the data for the monetary and frequency features after being divided by the lifetime feature. Similar to the previous approaches, additional visualization using a Sankey diagram can be seen in Figure 13.



(a) (b)

Figure 13. Sankey Diagram RFM/Lifetime1 and RFM/Lifetime2

From Figure 13a, it can be observed that consumers who are loyal to the "Make Up" category are mostly loyal and potential in the general category as well. Similarly, consumers who are loyal to the "Face Care" category are almost all loyal in the general category. Additionally, from Figure 13b, it is evident that consumers who are loyal to the "Make Up" category mostly do not purchase products from the "Face Care" category. On the other hand, consumers who are loyal to the "Face Care" category mostly do not purchase products from the "Make Up" category. This is consistent with the findings of the previous approaches.

3.5 Handling Outlier Data

Continuing from the data cleaning at the beginning, the discarded outlier data was collected into a new dataframe for each category in each approach. This outlier data was manually assigned to the existing segments, taking into account the characteristics of the segments that had already been formed.

3.6 Actions that can be taken

After determining the labels of all clusters from the three approaches, the majority of cluster labels include new customers, almost lost customers, lost customers, loyal customers, promising customers, and hibernate customers. In line with the goal of customer segmentation, which is to implement targeted marketing for each segment, there are actions that can be taken for segments with these cluster labels, including:

a. New Customer: Start building relationships, provide early success, and offer on-boarding support.
b. Almost Lost Customer: Provide recommendations for their last purchased items and offer limited time offers.
c. Lost Customer: Regain their interest through targeted campaigns or simply ignore.
d. Loyal Customer: Engage customers by seeking product reviews and upselling opportunities.
e. Promising Customer: Offer free trials and increase brand awareness.
f. Hibernate Customer: Present relevant brands to previous purchases and offer special discounts.

4. Conclusion

Based on the consumer segmentation conducted for Emina product consumers from PT Paragon Technology and Innovation, the following conclusions can be drawn:

a. Consumer segmentation using each approach is not sufficient and cannot be combined into one. There is a need to separate the categories of products, especially for Make Up, Face Care, and Others.
b. Each approach has its own consumer segmentation with the aforementioned product categories.

c. The RFM+Lifetime approach proves to be the most suitable for this consumer segmentation, as the vanilla RFM and RFM/Lifetime approaches were less relevant when applied to the consumer data.
d. The consumer segmentation using the RFM+Lifetime approach is divided into four categories: Make Up, Face Care, Others, and General, where Make Up is further divided into five segments, while the other categories are divided into four segments each.

5. Acknowledgement

Thank you to Universitas Multimedia Nusantara in Indonesia for providing a space for academics to conduct this journal research. Hopefully, this research will contribute significantly to the growth of technology in Indonesia

References

[1] M. R. Sofi, I. Bashir, M. A. Parry, and A. Dar, "The effect of customer relationship management (CRM) dimensions on hotel customer's satisfaction in Kashmir," *International Journal of Tourism Cities*, vol. 6, no. 3, pp. 601–620, 2020, doi: 10.1108/IJTC-06-2019-0075.

[2] A. Lubis, R. Dalimunthe, Y. Absah, and B. K. Fawzeea, "The Influence of Customer Relationship Management (CRM) Indicators on Customer Loyalty of Sharia Based Banking System," *GATR Journal of Management and Marketing Review*, vol. 5, no. 1, pp. 84–92, 2020, doi: 10.35609/jmmr.2020.5.1(8).

[3] D. Ngelyaratan and D. Soediantono, "Customer Relationship Management (CRM) and Recommendation for Implementation in the Defense Industry : A Literature Review Customer Relationship Management (CRM) dan Usulan Penerapannya Pada Industri Pertahanan : A Literature Review," *Journal of Industrial Engineering and Management Research*, vol. 3, no. 3, pp. 17–34, 2022, doi: 10.7777/jiemar.v3i3.279.

[4] R. Thakur and L. Workman, "Customer portfolio management (CPM) for improved customer relationship management (CRM): Are your customers platinum, gold, silver, or bronze?," *Journal of Business Research*, vol. 69, no. 10, pp. 4095–4102, 2016, doi: 10.1016/j.jbusres.2016.03.042.

[5] A. Naim, "Role of Information Systems in Customer Relationship Management," *International Journal of Intelligent Communication, Computing, and Networks*, vol. 02, no. 03, pp. 34–45, 2021, doi: 10.51735/ijiccn/001/37.

[6] A. A. Salameh, A. Hatamleh, M. S. Azim, and A. G. Kanaan, "Customer oriented determinants of e-crm success factors," *Uncertain Supply Chain Management*, vol. 8, no. 4, pp. 713–720, 2020, doi: 10.5267/j.uscm.2020.8.001.

[7] S. I. Wu and C. L. Lu, "The relationship between CRM, RM, and business performance: A study of the hotel industry in Taiwan," *International Journal of Hospitality Management*, vol. 31, no. 1, pp. 276–285, 2012, doi: 10.1016/j.ijhm.2011.06.012.

[8] S. Hansun, "FX forecasting using B-WEMA: Variant of Brown's Double Exponential Smoothing," *2016 International Conference on Informatics and Computing, ICIC 2016*, no. Icic, pp. 262–266, 2017, doi: 10.1109/IAC.2016.7905726.

[9] J. Wu *et al.*, "An Empirical Study on Customer Segmentation by Purchase Behaviors Using a RFM Model and K -Means Algorithm," *Mathematical Problems in Engineering*, vol. 2020, no. November 2017, 2020, doi: 10.1155/2020/8884227.

[10] O. Doğan, E. Ayçin, and Z. A. Bulut, "Customer Segmentation by Using RFM Model and Clustering Methods: A Case Study in Retail Industry," *International Journal of Contemporary Economics and Administrative Sciences*, vol. 8, no. 1, pp. 1–19, 2018, [Online]. Available: www.ijceas.com

[11] J. Wei, S. Lin, and H. Wu, "A review of the application of RFM model," *African Journal of Business Management*, vol. 4, no. 19, pp. 4199–4206, 2010.

[12] S. Hansun, A. Wicaksana, and M. B. Kristanda, "Prediction of jakarta city air quality index: Modified double exponential smoothing approaches," *International Journal of Innovative Computing, Information and Control*, vol. 17, no. 4, pp. 1363–1371, 2021, doi: 10.24507/ijicic.17.04.1363.

[13] S. Monalisa, P. Nadya, and R. Novita, "Analysis for customer lifetime value categorization with RFM model," *Procedia Computer Science*, vol. 161, pp. 834–840, 2019, doi: 10.1016/j.procs.2019.11.190.

[14] D. Chen, S. L. Sain, and K. Guo, "Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining," *Journal of Database Marketing and Customer Strategy Management*, vol. 19, no. 3, pp. 197–208, 2012, doi: 10.1057/dbm.2012.17.

[15] K. P. Sinaga and M. S. Yang, "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/ACCESS.2020.2988796.

[16] E. Hot and V. Popović-Bugarin, "Soil data clustering by using K-means and fuzzy K-means algorithm," *Telfor Journal*, vol. 8, no. 1, pp. 56–61, 2016, doi: 10.5937/telfor1601056H.

SAGA: Journal of Technology and Information Systems
Vol 1, Issue 3, August 2023, Page 79-89
ISSN: 2985-8933 (Media Online)
DOI: 10.58905/SAGA.vol1i3.171

[17] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," *Electronics (Switzerland)*, vol. 9, no. 8, pp. 1–12, 2020, doi: 10.3390/electronics9081295.

[18] R. V. Singh and M. P. S. Bhatia, "Data clustering with modified K-means algorithm," *International Conference on Recent Trends in Information Technology, ICRTIT 2011*, pp. 717–721, 2011, doi: 10.1109/ICRTIT.2011.5972376.

[19] E. Aytaç, "Unsupervised learning approach in defining the similarity of catchments: Hydrological response unit based k-means clustering, a demonstration on Western Black Sea Region of Turkey," *International Soil and Water Conservation Research*, vol. 8, no. 3, pp. 321–331, 2020, doi: 10.1016/j.iswcr.2020.05.002.