

SEGMENTATION OF OUTLETS WITH K-MEANS CLUSTERING BASED ON CUSTOMER TRANSACTION DATA TO DEFINE MARKETING STRATEGIES (CASE STUDY: GS SUPERMARKET)

Ayu Ghandhi Putri M¹, Abdul Fattah², Septiadi Nugroho³, Jerry Heikal⁴
gandhiayu@gmail.com¹, abdul.fattah82@gmail.com², nugrohoseptiadi@gmail.com³,
jerry.heikal@bakrie.ac.id⁴
Master of Management, Bakrie University

Abstract

The largest retail company in South Korea which expanded to Indonesia in 2016 by using the brand GS Supermarket. Currently growing and developing to consist of 7 branch outlets spread across the JABODETABEK area. To be able to survive, develop & compete in this line of business, GS Supermarket must have a marketing strategy that is right on target. In this study the author used the clustering method with the k-means algorithm. This method is done by grouping customers from all GS Supermarket branch outlets based on customer transaction data (total transactions, phone providers and payment method), The result from this study shows that there are 3 clusters consisting of Clusters Mythic which is formed with a high transaction category, using payment methods through EDC Mandiri and BCA, is dominated by consumers who use Telkomsel and XL providers. Epic cluster is formed from transactions below the average of other clusters, using cash payment methods and dominated by non-member consumers. Cluster Legend is formed from medium transactions even though the value is still below average and dominated by consumers with payment methods through EDC BNI. The results of this study showed that the total amount of shopping transactions was dominated by consumers using Telkomsel and XL providers and using payment methods through EDC Mandiri and BCA. So that, the recommended marketing mix strategy can be in the form of establishing certain cooperation and promotions with mobile phone providers and banks.

Kata Kunci: *K-Means, Clustering, Retail Business, Marketing Strategy.*

1. INTRODUCTION

The Covid-19 pandemic which hit almost all countries, not only caused health problems but triggered an economic recession. Various industrial sectors cannot carry out activities normally, such as the retail industry which is one of the business sectors that was affected by the Covid-19 pandemic. Some of retail companies have losses, some of them even closed their businesses. In addition to declining market demand, the challenges also come from the components of fixed expenses that must be paid by entrepreneurs. Large-scale social restrictions (PSBB) issued by the government to reduce the rate of spread of COVID-19 made business actors, especially retail stores, suffer from a significant decrease in their turnover or income. However, the pandemic that is suppressing the retail industry must be addressed quickly by business actors in order to adapt and innovate to improve production efficiency and offer the best products for consumers.

As one of the new retail businesses in Indonesia, GS Retail also felt this situation. The largest retail company in South Korea which expanded to Indonesia in 2016 by using the brand GS Supermarket. Currently growing and developing to consist of 7 branch outlets spread across the JABODETABEK area. By targeting the upper middle market segment, one of the attractions of GS Supermarket is that offers a variety of products such as fresh food, processed food, household needs, beauty products, cleaning products, electronic goods, and other imported products which directly from South Korea and claims to sell its products at lower prices.

To be able to survive, develop, and win the competition, PT. GS Retail Indonesia must be good at analyzing current market situation and the market situation in the future. Customer-oriented marketing strategy has an important role in managing good relationships with customers. For the marketing strategy to be right on target, customer segmentation can be used to group customers based on the same characteristics.

2. RESEARCH METHODS

The type of research used in this research is the type of applied research. Applied research is research to produce something that can be directly applied to solve a problem. In this study, the authors applied the k-means algorithm to the SPSS application program to classify the types of sales transactions with categories of high transactions, medium transactions and below-average transactions based on customer transaction data (total transactions, phone provider and payment method) in all GS Supermarket outlet by using the k-means method. The research data was taken from the object of research, namely GS Supermarket.

3. RESEARCH RESULTS AND DISCUSSION

1. Output K-Means Cluster

The formation of the k-means cluster in this study was carried out using the SPSS application program to describe the classification results obtained. The data is classified into 3 clusters and based on customer transaction data (total transactions, phone provider and payment method) at all GS Supermarket branch outlets, 3 cluster membership results are obtained: cluster 1 consists of 5 branch outlets (Cipondoh, Jati Asih, Jatibening, Kemang Pratama and Tourism Legend), cluster 2 consists of 1 branch outlet (Tajur), and cluster 3 consists of 1 branch outlet (Mampang).

Table 1 : Output final cluster centers (A)

	Cluster		
	1	2	3
Zscore(No_Number)	-.34408	2.16881	-.44843
Zscore(Cash)	.10635	.66470	-1.19646
Zscore(EDC_BCA)	.47453	-.84242	-1.53025
Zscore(EDC_BNI)	.32537	1.75320	-.12636
Zscore(EDC_MDR)	.54350	1.20836	-1.50915
Zscore(E_PAYMENT)	.28578	1.29623	-.13265
Zscore: Total Belanja	.54059	1.66594	-1.03703
Zscore(Indosat)	.37841	-.67740	-1.21464
Zscore(Telkomsel)	.57375	1.58603	-1.28270
Zscore(XL)	.48764	-1.12697	-1.31122

Table 1 shows that the final output of the cluster centers is still related to the previous data standardization process, which refers to the z-score with the following conditions:

- A negative value (-) means the data is below the average
- A positive value (+) means that the data is above the total average

In cluster 1 listed in table 1, it shows the GS Supermarket outlets that have the number of shopping transactions, the types of payments with various payment methods and the number of mobile phone users from various providers which is more than the average population of the outlets researched. This is evident from the positive (+) values contained in table 1 in all the variables.

Thus, it can be assumed that cluster 1 is a grouping of outlets visited by many consumers. Consumer data using Telkomsel's mobile phone provider and payment methods using Bank Mandiri's EDC machines dominate in cluster 1.

Table 2 : Output final cluster centers (B)

	Cluster		
	1	2	3
Zscore(No_Number)	-0.34408	2.16881	-0.44843
Zscore(Cash)	0.10635	0.66470	-1.19646
Zscore(EDC_BCA)	0.47453	-0.84242	-1.53025
Zscore(EDC_BNI)	0.32537	-1.75320	0.12636
Zscore(EDC_MDR)	0.54350	-1.20836	-1.50915
Zscore(E_PAYMENT)	0.28578	-1.29623	-1.13265
Zscore: Total Belanja	0.54059	-1.66594	-1.03703
Zscore(Indosat)	0.37841	-0.67740	-1.21464
Zscore(Telkomsel)	0.57375	-1.58603	-1.28270
Zscore(XL)	0.48764	-1.12697	-1.31122

In cluster 2, which is listed in table 2, it shows GS Supermarket outlets where almost all the assessment instruments are in a position below the average outlet population. This is indicated by a negative value (-), except for the cash payment method instrument and membership with an invalid provider number which is still above the average outlet population. So, it can be assumed that outlets with below average transactions are in cluster 2.

In cluster 3, which is listed in table 2, it shows GS Supermarket outlets where almost all the assessment instruments are in a position below the average population of outlets. This is indicated by a negative value (-), except for the payment method instrument using the BNI EDC machine. So, it can be assumed that outlets with medium transactions are in cluster 3.

Table 3 : Cluster membership

Cluster Membership			
Case Number	Location Name	Cluster	Distance
1	CIPONDOH	1	2.053
2	JATI ASIH	1	1.443
3	JATI BENING	1	2.202
4	KEMANG PRATAMA	1	1.375
5	LEGENDA WISATA	1	2.544
6	MAMPANG	3	.000
7	TAJUR	2	.000

Table 3 shows the results of the classification of GS Supermarket outlets based on customer transaction data (total transactions, phone provider and payment method). Three cluster membership results were obtained, namely: cluster 1 consisting of 5 branch outlets (Cipondoh, Jati Asih, Jatibening, Kemang Pratama and Legenda Wisata), cluster 2 consisting of 1 branch outlet (Tajur), and cluster 3 consisting of 1 branch outlet (Mampang).

From the classification results, it can be concluded that cluster 1 which consists of 5 branch outlets (Cipondoh, Jati Asih, Jatibening, Kemang Pratama and Legenda Wisata) can be categorized as an outlet that has high transactions, cluster 2 which consists of 1 branch outlet (Tajur). can be categorized as outlets with medium transactions and cluster 3 consisting of 1 branch outlet (Mampang) can be categorized as outlets with below average transactions.

To ensure differences between clusters, it can be done with ANOVA as shown in table 3. Based on these results it is known that each cluster for all the variables involved has a significant difference because the sig value obtained is less than 0.05.

Table 4 : ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore(No_Number)	2.748	2	.126	4	21.847	.007
Zscore(Cash)	.965	2	1.018	4	.948	.460
Zscore(EDC_BCA)	2.089	2	.456	4	4.583	.092
Zscore(EDC_BNI)	1.809	2	.595	4	3.040	.157
Zscore(EDC_MDR)	2.607	2	.196	4	13.279	.017
Zscore(E_PAYMENT)	1.053	2	.973	4	1.082	.421
Zscore: Total Belanja	2.656	2	.172	4	15.442	.013
Zscore(Indosat)	1.325	2	.837	4	1.582	.312
Zscore(Telkomsel)	2.903	2	.048	4	60.085	.001
Zscore(XL)	2.089	2	.455	4	4.587	.092

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

4. CONCLUSION

Based on the results of this study related to the grouping of customers from all GS Supermarket branch outlets based on customer transaction data (total transactions, phone provider and payment method), it can be concluded that:

Based on the k-means cluster output, 3 results of cluster membership are obtained, from the classification results it can be concluded that cluster 1 which consists of 5 branch outlets (Cipondoh, Jati Asih, Jatibening, Kemang Pratama and

Legenda Wisata) can be categorized as an outlet which has high transactions, cluster 2 which consists of 1 branch outlet (Tajur) can be categorized as an outlet that has below average transactions and cluster 3 which consists of 1 branch outlet (Mampang) can be categorized as medium outlet that has transactions.

REFERENCE

- Alma, B., & Hurriyati, R. (2008). *Manajemen Corporate dan strategi pemasaran jasa pendidikan*. Bandung: Alfabeta.
- Berry, M. A. (2000). Mastering Data Mining: The Art and Science of Customer Relationship Management. *Industrial Management & Data Systems*, 100(5). <https://doi.org/10.1108/imds.2000.100.5.245.2>
- Kotler, P., & Armstrong, G. (2006). *Principles of Marketing 14 E*. In Pearson Prentice Hall.
- Kotler, P., Saunders, J., Armstrong, G., & Wong, V. (2012). *Principles of Marketing Second European Edition*. In Erlangga (Vol. 66).
- Madhulatha, T. S. (2012). AN OVERVIEW ON CLUSTERING METHODS. *IOSR Journal of Engineering*, 02(04). <https://doi.org/10.9790/3021-0204719725>
- Wu, J. (2012). *Advances in K-means Clustering: a data mining thinking*. In Springer Theses: recognizing outstanding Ph.D. Research.