

ONLINE-CUSTOMERS SEGMENTATION ON PT. RUMAH MEBEL NUSANTARA (IKEA) IN SEMARANG REGION BY USING K-MEANS ANALYSIS WITH SPSS

Astried Finnia Ayu Kirana, Neisya Esvandiary Iswari, Damar Mafatir Romadhon, Jerry Heikal

Program Studi Magister Manajemen, Fakultas Ekonomi dan Ilmu Sosial Universitas Bakrie

Email: neisya.diary@gmail.com, astriedkirana@gmail.com, damarmafatier@gmail.com

Abstract

Business is an activity carried out by individuals and organizations that create value through goods and services to obtain a profit. In running a business, marketing strategy is the main key in product development for the progress of the company. Marketing is not only for product and service development, but also needs to take into account customer segmentation. Segmentation has done using K-Means Clustering based on SPSS. In this study, researchers took a case study at PT. Rumah Mebel Nusantara (IKEA) in the Semarang area by conducting research on 306 respondents to IKEA online customers. In this study obtained 3 clusters. Where cluster 1 is a Methodical Customer, Cluster 2 is referred to as a Humanistic Customer, and Cluster 3 is referred to as a Spontaneous Customer. Based on the results of data processing using SPSS, researchers chose Cluster 3 because this cluster has a high potential to increase sales with the type of purchase and the number of transactions are greater than the other 2 clusters. This cluster will tend to be more loyal than the other 2 clusters because this type of customer if they satisfied with the brand they will buy it.

Keywords: Online-Customer Segmentation, K-Means Clustering

Introduction

Business is an activity carried out by individuals and organizations that create value through goods and services to obtain a profit. In running a business, marketing strategy is the main key in product development for the progress of the company. Marketing is not only for product and service development, but also needs to take into account customer segmentation (Webster Jr, 1988). Where segmentation according to (Armstrong et al., 2014) is dividing a market into groups of buyers with different desires,

How to cite:	Astried Finnia Ayu Kirana, Neisya Esvandiary Iswari, Damar Mafatir Romadhon, Jerry Heikal (2023), Online-Customers Segmentation On PT. Rumah Mebel Nusantara (Ikea) In Semarang Region By Using K-Means Analysis With Spss, Vol. 8, No. 2, Februari 2023, http://dx.doi.org/10.36418/syntax-literate.v8i2.11378
---------------------	---

E-ISSN:	2548-1398
----------------	-----------

Published by:	Ridwan Institute
----------------------	------------------

characteristics or behaviors. After doing the segmentation, it will make it easier for the company to carry out the marketing process, where the process will be more focused because each segment already has similarities, both in terms of needs and behavior.

Grouping is done by using data mining. In data mining there is the term clustering, where this clustering has an important role which will divide the data into several clusters according to their similarities. K-Means clustering is a non-hierarchical data clustering method that groups data in the form of one or more clusters/groups (Khairani & Sutoyo, 2020). From this grouping will produce the characteristics of each customer in each cluster which can be used as a reference for determining marketing strategies.

K-Means Alorythm Classification

K-Means clustering is a non-hierarchical data clustering method that groups data in the form of one or more clusters/groups. Data that has the same characteristics are grouped in one cluster/group and data that has different characteristics are grouped with other clusters/groups so that data that is in one cluster/group has a small level of variation (Premana et al., 2017) .

According to (Bock, 2007) there are Clustering steps using K-Means, as follows:

1. Select the number of clusters k
2. Initialization of k cluster centers can be done in various ways. But the most often done is in a random way. Cluster centers are assigned initial values with random numbers.
3. Allocate all data/objects to the nearest cluster. The proximity of two objects is determined based on the distance between the two objects.
4. Recalculate the cluster center with the current cluster membership. Cluster center is the average of all data/objects in a particular cluster.
5. Reassign each object using the new cluster center. If the cluster center does not change again then the clustering process is complete.

Clustering

Clustering is the formal study of methods and algorithms for partitioning or clustering. Clustering analysis does not use the previous category labeling. Clustering is unsupervised learning or does not have a data training stage, in contrast to classification (Namdev et al., 2015).

Clustering is a method for finding and grouping data that have similar characteristics (similarity) between one data and another (Aldino et al., 2021). Clustering is a data mining method that is unsupervised; implemented without any training and without a teacher and does not require an output target. In data mining there are two types of clustering methods used in grouping data: hierarchical clustering & non-hierarchical clustering (Schonlau, 2004).

According to (Dib & Carbone, 2012) there are 2 types of clustering, namely Hierarchical Clustering & Non-hierarchical Clustering. Hierarchical clustering method is a method of grouping data that begins by grouping two or more objects that have the closest similarity. Then the process is passed to another object which has a second immediacy. And so on so that the cluster will form a kind of tree where there are 6 clear

hierarchies (levels) between objects, from the most similar to the least similar. Logically all objects in the end will only form a cluster. Dendograms are usually used to help clarify the hierarchical process (Almeida et al., 2007). Meanwhile, the non-hierarchical clustering method begins by determining in advance the desired number of clusters (two clusters, three clusters, or so on). After the number of clusters is known, then the cluster process is carried out without following the hierarchical process. This method is commonly known as K-Means Clustering (Jung et al., 2003).

Marketing Mix

(Pour et al., 2013) define the marketing mix as a set of tactical marketing tools that the company combines to produce the response it wants in the target market. (Melewar & Saunders, 2000) explains that there are 8 components in the marketing mix called the 8Ps, including:

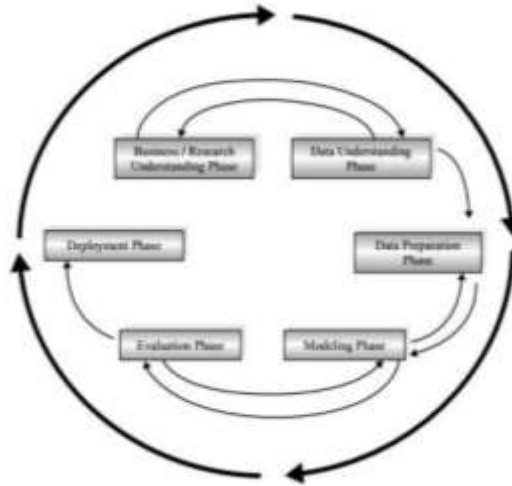
- 1) Product
It is a form of service organization offering aimed at achieving goals through satisfying customer needs and wants.
- 2) Price
It is the expenditure of money, time, and effort that customers sacrifice in buying and consuming the products and services that the company offers or provides.
- 3) People
Customers and employees involved in the activities of producing products and services (service production).
- 4) Place
It is a distribution decision regarding the ease of access to services for customers.
- 5) Promotion
It is all communication and incentive design activities to build the desired customer perception of the company for the specific services that the company provides.
- 6) Packaging
The way a company combines its main product or main service with additional products and services and makes it one price
- 7) Programming
Special events, programs, or promos for a certain period of time that have been determined to increase customer purchase spending or to increase attractiveness
- 8) Partnership
Promotional or marketing activities carried out cooperatively by tourism companies and organizations, or other companies and are mutually beneficial.

In this study, researchers took a case study at PT. Rumah Mebel Nusantara (IKEA), where IKEA is a retail company that focuses on online sales. PT. Rumah Mebel Nusantara (IKEA) has several delivery options, such as home delivery and delivery to business units in Semarang region. Based on this explanation about IKEA, the researchers are interested in creating and analyzing clusters of customers who buy products online at PT. Rumah Mebel Nusantara (IKEA) in the Semarang region.

Research Method

The research method used in this research is CRISP-DM (Cross Standard Industry for Data Mining). CRISP-DM was developed in 1996 by analysis from several industries such as standardization Daimler Chrysler (Daimler-Benz), SPSS, NCR (Nava & Hernández, 2012).

Gambar 1
Proses Data Mining CRISP-DM (Larose & Larose, 2014)



According to (Larose & Larose, 2014) the data mining process based on CRISP-DM consists of six phases, including:

1. *Business Understanding Phase*
 - a. Determination of project objectives and needs in detail within the scope of the business or research unit as a whole.
 - b. Translating goals and constraints into formulas for data mining problems.
 - c. Prepare initial strategies to achieve goals.
2. *Data Understanding Phase*
 - a. Collecting data.
 - b. Using data investigation analysis to identify further data and search for initial knowledge.
 - c. Evaluating data quality.
 - d. If necessary, select a small group of data that may contain the pattern of the problem.
3. *Data Preparation Phase*
 - a. Prepare from initial data, collect data that will be used for the entire next phase. This phase is heavy work that needs to be carried out intensively.
 - b. Select the cases and variables to be analyzed and the appropriate analysis to be carried out.
 - c. Make changes to some variables if needed.
 - d. Prepare the initial data so that it is ready for modeling tools.
4. *Modelling Phase*

- a. Select and apply the appropriate modeling technique.
 - b. Calibrate model rules to optimize results.
 - c. It should be noted that several techniques may be used for the same data mining problem.
 - d. If needed, the process can return to the data processing phase into a form that is in accordance with the specifications for the needs of certain data mining techniques.
5. *Evaluation Phase*
- a. Evaluate one or more models used in the modeling phase to obtain quality and effectiveness before being deployed for use.
 - b. Determine whether there is a model that meets the objectives in the initial phase.
 - c. Determine whether there are important business or research issues that have not been handled properly.
 - d. Taking decisions related to the use of the results of data mining.
6. *Deployment Phase*
- a. Using the resulting model. The formation of the model does not indicate that the project has been completed.
 - b. A simple example of deployment: report generation.

Result and Discussion

Based on the research results from 306 data processed, 3 clusters were produced where cluster 1 consisted of 199 customers, cluster 2 consisted of 63 customers, and cluster 3 consisted of 44 customers. The results of data processing can be seen in Table 4.3.

Tabel 1
K-Means
Number of Cases in each
Cluster

Cluster	1	2	3
Valid	199.000	63.000	44.000
Missing			.000

Meanwhile, the processing and classification is carried out on several variables including total transaction, normal price variable, discount price variable, customer address which is divided into city and regency, type of delivery consisting of home delivery and business unit, payment method consisting of debit, credit and E-Wallet, as well as order source websites and applications. Furthermore, from the results of data processing using SPSS obtained results such as table XX. From the 3 clusters, it is formulated into 3 personas which include:

1. Cluster 1 is referred to as a *methodical customer*, namely customers in this category are more inclined to compare various aspects, including the uses and needs of their

customers before buying. So that from the data under cluster 1, they spend more according to their needs, where in this case study the transaction value is the smallest compared to other clusters. In accordance with the character of the persona that customers use debit more than credit. As for the type of delivery, customers are more inclined to consider which method is more profitable, so that the value between home delivery and business units is almost equal in value even though the use of home delivery is still greater. Because customers with these characteristics prefer to compare before buying, customers in this persona are not too affected by discounts.

2. Cluster 2 is referred to as *humanistic customer*, that is, customers in this category are more inclined to put their emotions forward when buying something. From the results of data processing, it is found that this persona characteristic uses credit more than debit and uses business units as the delivery option. This means that the customer of this model prefers free shipping because it is taken at the business unit rather than at home. Meanwhile, the transaction value is still below the average of all transactions although it is larger than the cluster type 1. This group of customers all live in the city.
3. Cluster 3 is referred to as *spontaneous customer*, that is, customers in this category are more influenced and influenced by external customer factors. Another characteristic is that in this persona, customers who have bought a product and liked the product then influence and are influenced by other people at a certain time. Thus, from the processing results, it is found that customers with this characteristic are described by the use of applications in making purchases that are more dominant than other clusters. Customers who use the application are usually customers who believe in the product and buy the product repeatedly, because in getting the application the customer needs to download it on their smartphone. Another characteristic is also described that customers in this cluster use more discounts than others and the transactions carried out are the highest compared to others.

Table 2
Final Cluster Centers

	Cluster		
	1.00	2.00	3.00
Zscore(Total_Transaction)	-.39	-.17	1.14
Normal	.99	.95	.91
Discount	.01	.05	.09
City	.96	1.00	.98
Regency	.04	.00	.02
Home_Delivery	.56	.00	.41
Unit_Business	.44	1.00	.59
Debit	.79	.00	.73
Credit	.15	.83	.23
E_Wallet	.06	.17	.05

Website	.43	.38	.18
Apps	.57	.62	.82

Based on the results of the above processing, the researcher suggests choosing cluster 3 because this cluster has a high potential to increase sales with the type of purchase and the number of transactions is greater than the other 2 clusters. This cluster will tend to be more loyal than the other 2 clusters because this type of customer if he is satisfied with a brand he will buy it.

Therefore, it is necessary to do a marketing mix to increase the turnover of the company by using 8P's theory.

A. Product

Focusing to selling the accessories product (such as tablecloths, flower vases, tableware, etc) product with added value, so the added value can be marketed at price premium that can be leads to higher sales value & product strategy for growth, such as:

1. Volume strategy
 - a. Optimizing cost structure by determining aggressive pricing (compare with its class) in other competitors
 - b. Aggressive pricing can facilitate volume growth, especially applicable for commodities and /or mass market products
2. Innovation Strategy
 - a. Continues flow of improve and new product
 - b. Significant investment if necessary
 - c. Creative corporate culture is required
 - d. Fast implementation
3. Me too strategy
 - a. Benchmarks of successful competitive innovation
 - b. Small path between copy and differentiation
 - c. Flexibility and speed are very important

B. Promotion

1. Strategies that can be done are by:
 - a. Collaborating with companies providing delivery services and product sales media such as Go-To to to increase Aware interest Desire Action as IKEA said AIDA (Car Sticker Branding)
 - b. Conducting massive promotions (Geo-targetting) for sales areas in this study)
2. Key success factors in communication management
 - a. Get attention to the target product (Inquiry of the product)
 - b. Gain interest for products in the target group
 - c. Generate a desire in the target group to buy a product
 - d. Invite the target group to act in order to buy the product effectively
 - e. Manage all eight P's effectively and push each P to the best performance
 - f. Integrate all 8Ps into one marketing mix.

C. Partnership

- a. Adding partners as a pick-up point (Indomaret point) to increase sales
- b. Adding a delivery network of vendors to be more able to offer cheaper delivery services so as to increase the desire to buy goods.

D. Price

Communicates that the price offered is the best price compared to competitors with the same specifications and types.

E. Place

- a. Adding a free pick-up location in the area around Semarang.
- b. Adding a distribution network for pick-up areas such as indomaret points in the area

F. People

Providing training to employees to increase customer satisfaction, especially for after sales (presence, close and care to customers).

G. Process

Evaluate customer journey on the website and make improvements to the information provided to customers.

H. Physical Evidence

Ensure that each business unit performs the same process including the uniforms worn

Conclusion

Based on the results above, it can be concluded that there are 2 clusters of PT. XZY is cluster 1 (Methodical Customer): upper middle class community group with high transaction value, and cluster 2: middle class community group with low transaction value but large number of transactions. Companies are advised to target the second cluster by applying 8 Ps.

BIBLIOGRAPHY

- Aldino, A. A., Darwis, D., Prastowo, A. T., & Sujana, C. (2021). Implementation of K-means algorithm for clustering corn planting feasibility area in south lampung regency. *Journal of Physics: Conference Series*, 1751(1), 12038. <https://doi.org/10.1088/1742-6596/1751/1/012038>
- Almeida, J. A. S., Barbosa, L. M. S., Pais, A., & Formosinho, S. J. (2007). Improving hierarchical cluster analysis: A new method with outlier detection and automatic clustering. *Chemometrics and Intelligent Laboratory Systems*, 87(2), 208–217.
- Armstrong, G., Adam, S., Denize, S., & Kotler, P. (2014). *Principles of marketing*. Pearson Australia.
- Bock, H.-H. (2007). Clustering methods: a history of k-means algorithms. *Selected Contributions in Data Analysis and Classification*, 161–172. <https://doi.org/10.1007/978-3-540-73560-1>
- Dib, L., & Carbone, A. (2012). CLAG: an unsupervised non hierarchical clustering algorithm handling biological data. *BMC Bioinformatics*, 13, 1–14. <https://doi.org/10.1186/1471-2105-13-194>
- Jung, Y., Park, H., Du, D.-Z., & Drake, B. L. (2003). A decision criterion for the optimal number of clusters in hierarchical clustering. *Journal of Global Optimization*, 25(1), 91–111. <https://doi.org/10.1023/A:1021394316112>
- Khairani, N. A., & Sutoyo, E. (2020). Application of k-means clustering algorithm for determination of fire-prone areas utilizing hotspots in West Kalimantan Province. *Int. J. Adv. Data Inf. Syst*, 1(1), 9–16. <https://doi.org/10.25008/ijadis.v1i1.13>
- Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining* (Vol. 4). John Wiley & Sons.
- Melewar, T. C., & Saunders, J. (2000). Global corporate visual identity systems: using an extended marketing mix. *European Journal of Marketing*, 34(5/6), 538–550. <https://doi.org/10.1108/03090560010321910>
- Namdev, N., Agrawal, S., & Silkari, S. (2015). Recent advancement in machine learning based internet traffic classification. *Procedia Computer Science*, 60, 784–791. <https://doi.org/10.1016/j.procs.2015.08.238>
- Nava, J., & Hernández, P. (2012). Optimization of a hybrid methodology (CRISP-DM). In *Logistics Management and Optimization through Hybrid Artificial Intelligence Systems* (pp. 356–379). IGI Global. <https://doi.org/10.4018/978-1-4666-0297-7.ch014>

Online-Customers Segmentation on PT. Rumah Mebel Nusantara (Ikea) in Semarang
Region By Using K-Means Analysis With SPSS

Pour, B. S., Nazari, K., & Emami, M. (2013). The effect of marketing mix in attracting customers: Case study of Saderat Bank in Kermanshah Province. *African Journal of Business Management*, 7(34), 3272–3280. <https://doi.org/10.5897/AJBM12.127>

Premana, A., Wijaya, A. P., & Soeleman, M. A. (2017). Image segmentation using Gabor filter and K-means clustering method. *2017 International Seminar on Application for Technology of Information and Communication (ISEmantic)*, 95–99. <https://doi.org/10.1109/ISEMANTIC.2017.8251850>

Schonlau, M. (2004). Visualizing non-hierarchical and hierarchical cluster analyses with clustergrams. *Computational Statistics*, 19, 95–111. <https://doi.org/10.1007/BF02915278>

Webster Jr, F. E. (1988). The rediscovery of the marketing concept. *Business Horizons*, 31(3), 29–39. [https://doi.org/10.1016/0007-6813\(88\)90006-7](https://doi.org/10.1016/0007-6813(88)90006-7)

Copyright holder:

Astried Finnia Ayu Kirana, Neisy Esvandary Iswari, Damar Mafatir Romadhon,
Jerry Heikal (2023)

First publication right:

Syntax Literate: Jurnal Ilmiah Indonesia

This article is licensed under:

