

Analysis on the Application of Market segmentation

Jiayu Zhong^{*1}

¹Wheaton College, MA

*Corresponding author. Email: zhong_jiayu@wheatoncollege.edu

ABSTRACT

This paper introduces the role and benefits of market segmentation in general, and gives examples to elaborate in depth. It introduces some tools that can be used for segmentation, such as K-means, silhouette score, elbow method, etc. Then, this paper uses data from 300 people about personal consumption habits in E-commerce as an example to classify its clients. Finally, it analyzed a LEGO toy company's application of segmentation in their company.

Keywords: Market segmentation, K-means, Silhouette score, Elbow method, E-commerce

1. INTRODUCTION

Market segmentation, in a broad sense, is the process by which analysts study the consumption habits, needs, desires, and buying behavior of consumers (regular customers and potential consumers) to classify their company's overall consumer base into a number of different groups, then according to each group's differentiation, customers from each group should be targeted with different marketing methods. In other words, each consumer group is a small market, and the analyst needs to help the company understand the characteristics of each consumer group and customize specific recommendations for each group. In recent years, with the rise of e-commerce, big data has become an indispensable computing tool for large companies. Accordingly, market segmentation is an important work done by analysts using big data.

Customers, as an inevitable part of any market, are essential to be understood and utilized effectively and efficiently. To achieve that, we have to know what exactly is market segmentation? How does an analyst make segmentation? Most importantly, how can we apply market segmentation in real life, in industries? All above-mentioned questions would be answered with details in this paper. Specifically the last question, and also the key question of this paper, we would compare the market segmentation of some leading companies to come up with the most frequently used methods. In this paper, we will find the answer by analyzing some real data of the world-widely known toy company, Lego., which hopefully could make people have a deeper understanding of market segmentation.

2. THE FUNCTIONS OF MARKET SEGMENTATION

First of all, understanding the basic form of market segmentation would be a safe start to illustrate this paper. Generally speaking, there are four types of segmentation, the first one is demographic segmentation which can primarily describe basic profile data about a group of customers, such as, age, gender, income, education, etc. The second one is geographic segmentation, which mainly involves customers and clients' geographic location, such as ethnicity, country, city, etc. Noticeably, geographic information is seen as the original segmentation criterion used for the purpose of market segmentation [1]

Then, psychographic segmentation mainly includes the personal characteristics of customers and what would drive them to interest, namely outlook on values, interests, motivations, etc.. Haley[4] explains that the word "psychographics" was intended as an umbrella term to cover all measures of the mind. Finally, the last one is behavioral segmentation, which mainly focuses on how customers and clients behave, such as spending habits, purchasing habits, stickiness, etc.

With using marketing segmentation, a company could have a better understanding of how to more efficiently attract high-value customers and clients, in terms of their frequent spending habits and stickiness. Technically speaking, one of the most important purposes of a business is to attract customers to buy products. Furthermore, since after segmentation, the firm would at least have a brief sketch of their customer image, so the firm could adjust an optimal plan to throw in targeted

advertisement to not only the existing customers but also the potential ones. Moreover, by the known customers segmentation, this firm can make some customization for some valuable clients. In other words, some activities could be held to increase the customer stickiness. For example, for the “gold eagle group”, the business might hold some private, high-end events and special customized services, such as dedicated customer service. As many researches point out, the purpose of marketing is to match the genuine needs and desires of consumers with the offers of suppliers particularly suited to satisfy those needs and desires. [1]

3. METHODS ON SEGMENTATION

As discussed on the previous pages, segmentation plays a rather positive role in the contemporary market. Several common and popular methods about segmentation would be introduced next. First, one of the most common methods of segmentation nowadays is clustering. Clustering is a data mining technique used to analyse data that has variations and the number of lots. It is the process of grouping data into a cluster, by containing data that is as similar as possible yet distinguishable from other cluster objects.[3] There are mainly two kinds of clustering, hierarchical clustering analysis (HCA) and partitional clustering. The former one establishes a hierarchy of clustering, it could be either top-down or bottom-up process. The partitional clustering concept started with K-means algorithm was published in 1957.[5]) In this paper we mainly talked about the partitional clustering, specifically the category with K-means algorithm. In the K-means clustering, we should watch out the Z scale. The erroneous choice of Z

scale would make the whole clustering look strange. After we use K-means clustering, the Silhouette Score should be introduced to examine the “goodness” of a clustering. To be more precise about the “goodness”, it means the compactness and nearness of one group against the other data.

4. ANALYSIS ON THE DATABASE

For this part, it is the time to start applying what was mentioned in the previous context in this paper to real life. In recent years, machine learning is widely used to support human decision making in several related scientific areas such as medical, biological, image processing and so on. However, in market classification, marketing segmentation is facing numerous challenges in defining specific target customer zones to help enterprises launching their marketing strategies.[2] To show its power, we will analyze a real-life database about the annual revenue of about 300 consumers and how much they spend in e-commerce over the year, and take a look at the problems in this process and find ways to solve them. The reason why we choose this database is that it is not very large, does not involve too much redundant information, and is relatively easy to analyze. Firstly, we imported all the data into jupyter notebook and format a scatter plot. We can see that before we did any analysis and simplification of this database, the whole graph presents a very irregular and chaotic figure. Furthermore, when we look at the whole data in another 3-dimension version, the data is concentrated in a small part of the cube, which in another way indicates that present data did not perform satisfactory results.

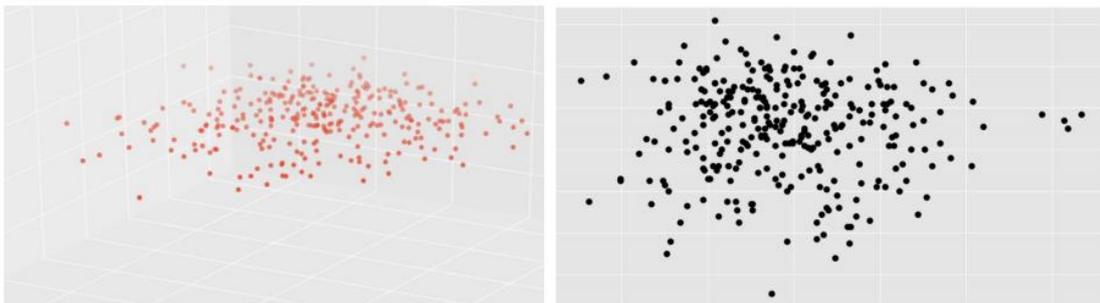


Figure 1 3-D diagram

I used elbow method and silhouette score to look for an optimal scheme that best fits this database. Unfortunately, the elbow method, also known as the knee of a curve, refers to the fact that we can determine our clustering only if we could find a clearly curved point in the graph.[8] That is, the graph in the notebook does not give us a precisely curved point or data, which means that we cannot use the elbow method to determine the clustering. We came up with the conclusion that normally when the silhouette score of this database is between 4 or

6, the silhouette score is maximized. Therefore, both of the two classification methods should be taken separately. Usually, researchers will use a combination of K-Means method with elbow method to improve efficiency of k-means performance in processing large amounts of data.[3] In this case, we only have 300 objects in total, which is not a large number in the marketing world. That is, clustering by k-means in this case might be inappropriate.



Figure 2 4 clustering



Figure 3 6 clustering

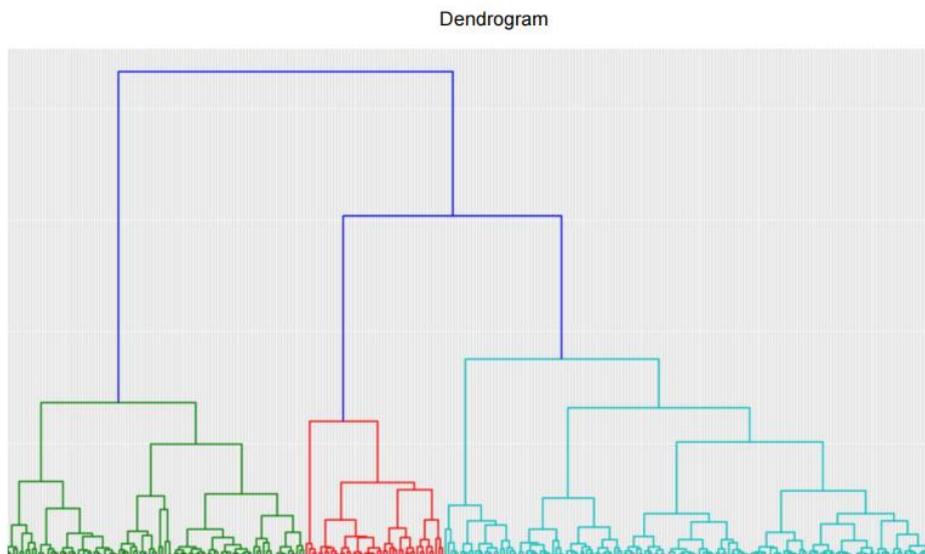


Figure 4 The tree-like graph

Then, I chose to use a dendrogram (a tree-like graph) to find whether or not it intrigues new ideas. As shown in graph (Figure 1), it is clear that the whole database can be divided into three categories represented by three colors. When we set the clustering equal to three, the whole set of data also looks effectively divided into three groups. Moreover, in this situation, their center points are all in the middle of each group, which indicates that the differences and distinctions between them are pulled apart.

After this complete analysis, we were not able to determine the "goodness" of clustering. Undoubtedly, the more comprehensive the data, the closer the ideal results we obtained would be to the truth. In this segmentation, main variables are the income of 300 people in one year and the expenditure in e-commerce in one year. Other personal details of these customers, such as geographic location, age, and ethnicity, are not specifically available, and these factors can make a difference in the results.

5. CASE STUDY: LEGO & ETC

In this part, LEGO would be mentioned as an example to examine how to classify customers (referred as Lego). As we all know, Lego is a world-class toy company, which originated in Denmark, with small building blocks of various colors put together to form thousands of different objects, as small as a little lamp, to a large aircraft carrier. In just a few decades, Lego has gone from a small, obscure company to a world-wide famous company. Lego's success is inseparable from its accurate portrayal of its target customers. Understanding consumer behavior in regards to its products gives the Lego Group an edge in developing social media strategies. In fact, it enables the toy maker to effectively use personas to segment its markets. [6] Moreover, according to Conny Kalcher, one of Lego's representatives, categorize Lego's customers and clients into six different groups, which are lead users, 1:1 community, connected community, active household, covered households, and all households. These six categories separately represent a group that is highly related to LEGO company. As I mentioned earlier, there are four types of segmentation, and each of the six categories is based on a holistic analysis of the customer, which not only looks at the consumer's spending habits and income, but also takes into account regional factors, age factors, ethnicity, and so on.

Another case study is about Electronic Toll Collection (ETC) in China. Electronic toll station, as a substitute for manual toll station, its appearance undoubtedly saves people time, and also saves the construction and operation cost of manual toll station. In this paper, we are mainly studying the classification of customers for ETC users with passing vehicles as the main research object. According to the experiment of Chao, Q[7], they did this by analyzing the entire sample

of ETC customers through the Cluster Large Application (CLARA) algorithm. In this case study, they also used decision tree to make it more clear. To be more specific, decision tree is composed of a leaf node and root nodes. Moreover, the final form of decision is a box-like graph. One of the advantages of a decision tree is its brief and easy-to-understand form, and it can cope with both numerical and categorical data. This decision tree contains six leaf nodes. The first line in each node displays the final fitted classification of observations (ETC customers), the second line shows the probability per classification[7] This kind of market segmentation indeed provides theoretical support for increasing the ETC customer scale and payment ratio, thereby improving the decision-making level of expressway operation and management.

6. CONCLUSION

This paper found that market segmentation indeed greatly help people to have a better understanding of their customers and purpose of their products. It used several methods to find out the best clustering of the database, which also proves that although it seems that we only need to classify and analyze databases, we actually use different methods for different research subjects. Different measurements can not only bring different results to make us reconsider the consequences, but also can give us more perspectives to think of the best solution. Moreover, in the example, we also encounter a problem that 300 people is a small database, although smaller database can make us easier to operating the customer segmentation in an experiment setting; however, in the real-world, we are suppose to have as many as possible data to make the results more accurate.

In the real-life example analysis, Lego performs a rather excellent role model for us. It's clear customer segmentation comprehensively takes into account all the details, from the age, economy and family situation of the client to the geographical location and buying habits of the client. Furthermore, the example of ETC also present us the effectiveness of customer segmentation.

ACKNOWLEDGMENT

First and foremost, I would like to show my deepest gratitude to teachers and professors in Wheaton college and in the business analytics project, who have provided me with valuable guidance in every stage of the writing of this thesis.

Further, I would like to thank my friends, Ritz, Wennie, Chole for their encouragement and support. Without all their enlightening instruction and impressive kindness, I could not have completed this paper.

REFERENCES

- [1] FRIEDRICH., D. S. (2019). MARKET SEGMENTATION ANALYSIS: Understanding it, doing it, and making it useful (1st ed.). Springer, Gapore: SPRINGER Verlag, SINGAPORE. doi:<https://doi.org/10.1007/978-981-10-8818-6>
- [2] Kieu, L. M., Ou, Y., & Cai, C. (2018). Large-scale transit market segmentation with spatial-behavioural features. *Transportation Research*. https://nls.ldls.org.uk/welcome.html?ark:/81055/vdc_100059619801.0x00001a.
- [3] Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018, April). Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In *IOP Conference Series: Materials Science and Engineering* (Vol. 336, No. 1, p. 012017). IOP Publishing.
- [4] Haley RI (1985) *Developing effective communications strategy – a benefit segmentation approach*. Wiley, New York
- [5] Nanda, S. J., & Panda, G. (2014). A survey on nature inspired metaheuristic algorithms for partitional clustering. *Swarm and Evolutionary computation*, 16, 1-18.
- [6] Barker, M. S., Barker, D., Bormann, N. F., Roberts, M. L., & Zahay, D. L. (2017). *Social media marketing : a strategic approach* (Second edition, student). Cengage Learning.
- [7] Chao, Q., Meng, Y., Peiqi, L., & Shuguang, L. (2018). Application of customer segmentation for electronic toll collection: a case study. *Journal of Advanced Transportation*, 2018. <https://doi.org/10.1155/2018/3635107>
- [8] Tynan AC, Drayton J (1987) Market segmentation. *J Mark Manag* 2(3):301–335