Segmentation of Students Using K-Means Clustering Case Study of Sekolah Dasar Tahfizh Quran (SDTQ) Citamulia

Aulia Firdaus[1*], Jerry Heikal[2]
[1,2]MM MBA Program, Bakrie University and Central Queensland University, Jakarta, Indonesia
*Corresponding author: aulia.firdaus2021@gmail.com

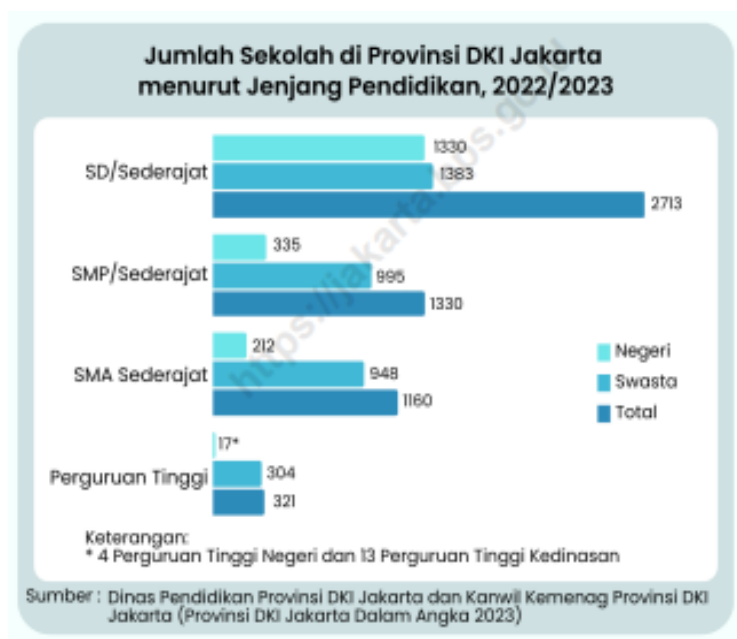| ARTICLE INFO	ABSTRACT
Keywords: k-means, clustering, persona, student profile, development policies	This study aims to determine consumer segmentation using K-Means Clustering in the case of SDTQ Cita Mulia elementary school students. The results of the study can be useful for preparing marketing strategies and sharpening the school's competitive advantages. The process consists of grouping consumer data based on several variables, which are student's gender, distance from home to school, age of student parents, parents' occupation and income, parents' education and the curriculum used by the previous school. After grouping the data, the initialization stage is carried out by giving an initial number. Then the clustering process is carried out. The resulting cluster will be analyzed to see the current Student profile. By studying the clusters formed, it helps us to determine school development policies, especially how marketing strategies can attract consumers in the middle cluster, namely those who have not joined us much but have the potential to be interested in our products. The result of the process there are four clusters with each cluster has specific persona name as follows: The first Cluster is the reachable distance with 5-6 km distance from the school, the second cluster in The long distance with 6 up to 28 Km distance from the school, the third cluster is the neighbor with just 3-4 km from school, and the forth cluster is the have this cluster persona has the highest parents income with the distance from school is 5-6 Km. The largest distribution of members is cluster 3 followed by cluster 4. Meanwhile, cluster 1 is not too big yet but still looks potential, and the lowest is cluster 2, so from this data the target persona is in cluster 3, the Neighbor persona.

INTRODUCTION

Basic education, especially at the age of 6 to 12 years, is considered a very important level of education in determining the success of a child's future. This is because at this age, a child is at the golden age in his growth and development, both physically, cognitively, character and emotionally. Cognitively, according to Piaget's theory, a child enters the concrete operational stage, which occurs at the age of 7 to 11 years. Cognitive development at this stage is marked by the development of organized and rational thinking. (Piaget, 1972). This concrete stage marks the beginning of logical

Aulia Firdaus[1*], Jerry Heikal[2]

thinking. Optimal cognitive development is important because cognitive function is the main supporter of other child developments, here they can understand the cause-and-effect process, understand the relationship between various ideas and are able to solve problems logically (Ginsburg & Opper, 1979).

Therefore, many parents are quite selective in determining the type of basic education for their children. This can be seen from the increasing growth of private elementary schools that are present to meet the needs of today's parents who are looking for quality basic education that cannot yet be fully obtained from existing public elementary schools.

Jumlah Sekolah di Provinsi DKI Jakarta menurut Jenjang Pendidikan, 2022/2023

Keterangan:
* 4 Perguruan Tinggi Negeri dan 13 Perguruan Tinggi Kedinasan

Sumber : Dinas Pendidikan Provinsi DKI Jakarta dan Kanwil Kemenag Provinsi DKI Jakarta (Provinsi DKI Jakarta Dalam Angka 2023)

Citamulia Islamic Elementary School was established in 2016 in South Jakarta. This school was initiated to answer the community's need for quality Islamic basic education. The growth of this school has increased significantly since entering the 2nd to 5th year of its establishment. However, after the pandemic year, namely 2022, the school experienced a decline in growth when viewed from the number of students. This condition requires evaluation and improvement in various fields, one of which is in the field of marketing.

In order to grow, schools must correctly determine the consumer segment that suits the products offered and study the right marketing strategy to expand the market or consumers.

By understanding consumer profiles, schools can conclude their largest consumer segments and then sharpen the quality of consumer needs in that segment. The second is, schools can also see other market potentials that are not yet large but have the potential to grow. By studying the profile of this segment, schools can determine the right marketing strategy.

The use of the K-Means Clustering Model with the SPSS program in this study is to segment customers. By using clustering methods such as K-Means, customer data can be grouped into clusters based on certain similar characteristics. (Jain, 2010)

This study will create clustering to determine consumer segmentation at the Citamulia Islamic Elementary School. After the cluster is formed, potential segments for schools can be identified and classified in order to increase their growth and determine the right marketing strategy by looking at the profile of the segment.

One of the unique value propositions of this school is preparing the student to memorize Al Quran from 5 to 15 Juz within 6 years elementary school period, beside the teaching of the national plus curriculum based on the Ministry of education standard for Elementary School in Indonesia.

To support students in memorizing the Quran at SDTQ Cita Mulia, several key elements has been implemented:

- Structured Memorization Program, manageable schedule for students to memorize specific portions verse daily.
- Qualified and experienced Quranic teachers with a deep understanding of tajweed (rules of Quranic recitation) that guide students not only in memorizing but also in ensuring correct pronunciation and understanding of the verses.
- Daily recitation program, regular recitation helps in reinforcing the memory, each students has their own personal Quran daily program, whether individually or in groups.
- Interactive Learning Environment, engage students in Quranic memorization through interactive and motivational activities. Group sessions, Quranic competitions, and rewarding progress in order to create positive learning environment.
- Parental involvement, parents play a crucial role in supporting Quranic memorization outside school. Providing resources and encouraging recitation at home strengthens the learning process.
- Balanced curriculum, while the primary focus is on Quranic memorization, students also need time for relaxation and recreational activities to avoid burnout, balancing their spiritual and academic growth.
- Peer support and motivation, encourage a community atmosphere where students motivate each other. Regularly celebrating progress, whether through certificates, events, or even simple acknowledgments in class, can drive motivation.

The school needs strategy to find the right students for school like SDTQ Cita Mulia, which focuses on both Quranic memorization and national-plus academic standards, so that it will make the learning process as ideal as targeted, steps of finding the right students are as follows:

Define clear criteria for student candidate, the school should target students who have a strong interest or background in Islamic school curriculum, since the school also teaches a comprehensive national-plus curriculum, academic readiness is also essential,

Aulia Firdaus[1*], Jerry Heikal[2]

assessing potential students through entry-level observation can help ensure they can follow the dynamic of the school.

Organize school events, hold open houses, Quranic competitions, and public events where parents can see the school's facilities, meet teachers, and understand the curriculum. These events showcase the school's unique value proposition and attract families looking for a balanced Islamic and academic education for their children.

Engage with Islamic communities by form partnerships with Islamic kindergarten schools, Quranic clubs, and organizations within the community. These groups often have students or families already inclined towards Islamic education and may refer students who are a good fit for SDTQ Cita Mulia. Offer scholarships for promising students from lower-income families who show potential for Quranic memorization and academic success.

Leverage word of mouth from Parent's alumni success stories, encourage to share their stories of successful alumni who have memorized significant portions of the Quran or who have excelled in academic or spiritual fields, invite testimonials and personal experiences are powerful tools for attracting like-minded families. Encourage current parents and students to spread the word about the school within their communities, referrals from trusted friends and family can be a strong influence on enrollment decisions.

Analyze the proper target market and create suitable marketing materials and program, with maintaining good offline and online presence. Ensure the school's website and social media accounts reflect its values and success stories. Post regular updates about students' Quranic achievements, competitions, and other school activities to attract the right audience. Targeted Advertising by running social media ads targeting Islamic families interested in Quranic education and holistic Islamic environments. Highlight the school's balance between Quranic memorization and academic excellence in your marketing materials.

METHOD

This study uses real data from the Cita Mulia school consumer database, namely student data in the 2023 and 2024 educational years. Consumers recorded in this study are consumers who are in two school growth conditions, namely in the year of increasing growth conditions and also consumers in the year of decreasing growth conditions. So that it is expected to provide a more relevant picture.

The stages of data processing in this study are as follows:

Data collection stage

This stage is the process of preparing raw data. Data is obtained directly from consumers or students who have entered the Cita Mulia Islamic School.

Data grouping stage

At this stage, the selection of attributes or variables that will be used in the data grouping process is carried out. The selection of this variable is based on relevant analysis needs.

The variables used to view consumer profiles are student gender, distance from the student's home to school, parent age, number of children owned, parent occupation and income and parent educational background and finally the curriculum used in the previous school.

The following is a division of group types based on each variable:

Table 1. Data Grouping Stage

No.	Student's Gender
1.	Male
2.	Female

No.	Parent's occupation
1.	Government employee
2.	Private employee
3.	Entrepreneur
4.	Others

No.	Parent's Income (Monthly)
1.	under 10 million
2.	between 10 to 20 million
3.	above 20 million

No.	Parent's Education
1.	Highschool Graduated
2.	Undergraduated
3.	Postgraduated

No.	Kindergarten's Curriculum
1.	Islamic
2.	National
3.	International
4.	Others

Data Processing Stage

At this stage, the selected variables are analyzed using the SPSS application using the K-Means method with clustering. The number of clusters determined is 4 clusters. This is because it can produce segments with clearer variations.

Results Analysis Stage

Aulia Firdaus[1*], Jerry Heikal[2]

Table 2. Initial Cluster Centers

	Cluster			
	1	2	3	4
Male	1	0	0	1
Female	0	1	1	0
Distance	13	33	1	3
Government Employee	0	0	1	0
Employee	1	1	0	1
Entrepreneur	0	0	0	0
Others	0	0	0	0
Child	1	2	3	1
Islam	1	0	1	0
National	0	0	0	0
International	0	0	0	0
Others	0	1	0	1
A	0	0	0	0
B	1	0	1	1
C	0	1	0	0
Father	56	39	51	30
Mother	52	36	37	30
SMU	1	0	0	0
S1	0	1	1	1
S2	0	0	0	0
SMU	0	0	0	1
S1	1	1	1	0
S2	0	0	0	0

At this stage, an analysis is carried out from the clustering results formed in the data processing process.

RESULT AND DISCUSSION

Table 3. Iteration History

			Change in Cluster Centers	
Iteration	1	2	3	4
1	8.159	5.828	8.952	9.301
2	2.376	0.000	0.495	0.842
3	0.409	0.000	0.179	0.147
4	0.193	0.000	0.062	0.000
5	0.000	0.000	0.000	0.000

After the clustering process using the k-means algorithm, four clusters were identified consisting of several members with similar characteristics. Information about the members in each cluster can be seen in Table 1, 2, 3 and Table 4.

Table 4. Cluster

	1	2	3	4
Male	0.54	0.80	0.52	0.60
Female	0.46	0.20	0.48	0.40
Distance	5.64	28.00	3.35	4.62
PNS	0.25	-	0.22	0.18
Employee	0.46	0.40	0.54	0.57
Entrepreneur	0.14	0.40	0.18	0.16
Others	0.14	0.20	0.05	0.09
Child	2.89	3.60	3.03	2.51
Islam	0.68	0.60	0.64	0.72
National	-	0.20	0.09	0.07
International	0.04	-	0.05	0.06
Others	0.29	0.20	0.22	0.15
A	-	-	0.05	0.11
B	0.61	0.20	0.61	0.62
C	0.39	0.80	0.34	0.27
Father	50.82	37.80	43.16	37.71
Mother	47.04	34.80	41.35	35.94
SMU	0.25	-	0.11	0.13
S1	0.64	1.00	0.71	0.65
S2	0.11	-	0.18	0.22
SMU	0.21	0.40	0.09	0.24
S1	0.75	0.40	0.85	0.63

Aulia Firdaus[1*], Jerry Heikal[2]

			0.07	0.12
S2				
	0.04	0.20		

Based on the results of this clustering, it can be seen that the current school consumer segmentation is divided into 4 clusters. The dominant variable that differentiates each cluster is the distance between home and school. The next is the age range of parents. While other variables, which are income, education, parents' jobs and previous school curriculum and the number of dependent children in the family, tend to be insignificant, which can be interpreted that the current consumer segment is quite homogeneous when viewed from these variables. Here are the details for each cluster:

Cluster 1 The Reachable Distance

Majority of students are male, distance from home to school is 5 to 6 km, the majority of parents' jobs are private employees with incomes between IDR 10 to 20 million monthly, educational background is undergraduate, choosing an Islamic curriculum during kindergarten.

Cluster 2 The Long Distance

Majority of students are male, distance from home to school is 28 km, the parents' jobs percentage share the same portion between private employee and entrepreneur with incomes between IDR 10 to 20 million monthly, educational background is undergraduate, choosing an Islamic curriculum during kindergarten.

Cluster 3 The Neighbors

Majority of students are male, distance from home to school is 3-4 km, the majority of parents' jobs are private employees with incomes between IDR 10 to 20 million monthly, educational background is undergraduate, choosing an Islamic curriculum during kindergarten.

Cluster 4 The Have

Majority of students are male, distance from home to school is 4-5 km, the majority of parents' jobs are private employee with incomes above IDR 20 million monthly, educational background is undergraduate, choosing Islamic curriculum during kindergarten.

If we try to analyze the distribution of the number of members in a cluster, as seen in table 4, it can be concluded that the largest consumers are cluster 3 followed by cluster 4. Meanwhile, cluster 1 is not too big yet but still looks potential, because it is reviewed from the similarity of the profile on other variables with the profiles of clusters 3 and 4.

Table 5. Number of Cases in each Cluster

Cluster	1	28.000
	2	5.000
	3	92.000
	4	82.000
Valid		207.000
Missing		0.000

The main target market that most suitable for SDTQ Cita Mulia is in the cluster 3 with The Neighbor persona, with the short distance and majority of parents' jobs are private employees with incomes between IDR 10 to 20 million monthly, SDTQ Cita Mulia should give more value for this target market, such as offer more admission cost discount for neighbor's criteria, or give more scholarship for student from surrounding area.

To achieve the goal of increasing school growth, there are two things that can be done. These are:

1. Strengthening the advantages of schools based on the reasons why parents currently choose Cita Mulia schools, which are the distance is not too far, the price is medium and the Islamic curriculum. The pricing policy should not move far from the standard medium price and continue to develop the Islamic curriculum, especially those that can be directly felt by students.

2. For market expansion purposes, the most potential cluster to target is cluster 1. Schools can adjust massive advertising movements to an area of about 6 km from the school and carry out promotional or educational content that is in accordance with the character or mindset of parents aged 47 to 51 years old.

3. In order to attract prospect students in cluster 2, schools must think about improving more complete and modern school facilities, as well as facilitating pick-up and drop-off services. This is because the significant difference between cluster 2 and the others is the variable of a further distance from home and a higher income of parents, of course they will pay close attention to facilities as a consideration.

4. Majority of student's gender are male. In order to increase the school's growth, the marketing strategies may give the more expose of female theme, for example the content of school's social media accounts can expose female student's activities in school, or make a simple fun or educative content about girl's subject.

CONCLUSION

Based on the clustering process of Cita Mulia elementary school student data in 2023-2024, it can be seen that Cita Mulia students are currently divided into 4 clusters. It can also be seen that consumer segmentation is still very homogeneous, which are those homes are about 3-5 km away, have middle incomes and like the Islamic curriculum as a school choice.

SDTQ Citamulia should offer more advantages to the target market where the distance is not too far is a must, with medium price and Integrated Quranic and national curriculum, combined with massive advertising In order to attract prospective students. Schools must think about improving school facilities, as well as facilitating pick-up and drop-off services. In order to increase the school's growth, the marketing strategies may give the more expose of female themes and educative contents.

Aulia Firdaus[1*], Jerry Heikal[2]

REFERENCES

Piaget, J., & Inhelder, B. (1972). *The psychology of the child* (H. Weaver, Trans.). Basic Books. (Original work published 1966)

Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letters, 31*(8), 651–666. https://doi.org/10.1016/j.patrec.2009.09.011

Ginsburg, H. P., & Opper, S. (1979). *Piaget's theory of intellectual development* (2nd ed.). Prentice-Hall.

Praditya, R. G., Sembodo, G., & Heikal, J. (2024). Market segmentation analysis to find out products

and services that suit customer needs using the python KMEANS clustering method (Case study: Superindo Tambun Area, Bekasi). *Jurnal Teknik Industri Terintegrasi (JUTIN), 7*(4), 2072–2081. https://doi.org/10.31004/jutin.v7i4.35889

Putri, J. H., & Heikal, J. (2024). Customer Segmentation Using Python K-Modes Clustering Method (Case Study on Netflix Userbase). *Innovative: Journal Of Social Science Research*, *4*(5), 1048–1059. https://doi.org/10.31004/innovative.v4i5.1488

Santosa, S., & Heikal, J. (2024). Analysis of Global Bank's Financial Performance with the Clustering K-Means Model. *JRAP (Jurnal Riset Akuntansi Dan Perpajakan)*, *11*(2), 283-288.

Saputra, T. C., Fadhilah, S. M., Mangkuto, S. U., & Heikal, J. . (2024). Segmentation, targeting and positioning analysis using k-means clustering model: A case study of the laptop market in Indonesia. *International Journal of Applied Finance and Business Studies*, *12*(2), 195–203. https://doi.org/10.35335/ijafibs.v12i2.313

Apriani, A., & Heikal, J. (2024). Segmentation Analysis Using K-Means Clustering Model with SPSS Case Study of Backpacker Jakarta Community Members. Jurnal Indonesia Sosial Sains, 5(03), 441–460. https://doi.org/10.59141/jiss.v5i03.1031

Ardiansyah, G. T., Santosa, S., Hasibuan, M. S., & Heikal, J. (2024). Mapping the Wuling vehicle market with K-Means Clustering : An effective digital marketing strategy. 14(2), 136–150.

Ariati, I., Norsa, R. N., Akhsan, L., & Heikal, J. (2023). Segmentasi Pelanggan Menggunakan K-Means Clustering Studi Kasus Pelanggan Uht Milk Greenfield. Cerdika: Jurnal Ilmiah Indonesia, 3(7), 729–743.

Farhan, M., & Heikal, J. (2024). Used Car Customer Segmentation Using K-Means Clustering Model With SPSS Program: Case Study Caroline.Id. Jurnal Indonesia Sosial Sains, 5(03), 543–559. https://doi.org/10.59141/jiss.v5i03.1042

Mulyo, I. A., & Heikal, J. (2022). Customer Clustering Using The K-Means Clustering Algorithm in Shopping Mall in Indonesia. Management Analysis Journal, 11(4), 365–371.

Waskita, R. M. A., Ferli, I., Fahrizal , R. R., & Heikal , J. (2024). Customer Segmentation Based On Age, Gender, Product And Total Customer Balance At Bank Xyz Using The K-Means Clustering Model. *Neraca: Jurnal Ekonomi, Manajemen Dan Akuntansi*, *2*(8), 550–561. Retrieved from https://jurnal.kolibi.org/index.php/neraca/article/view/2172

Ayu Pradina, D. ., Kurniawati, Y., Syawaldi Afwan, A. ., & Heikal, J. . (2024). RFM Segmentation and K–Means Clustering of Skincare Product (Case study

Scarlett). *Jurnal Sains Dan Teknologi*, *6*(2), 213-216. https://doi.org/10.55338/saintek.v6i2.3644

Perdhana, R., & Heikal, J. (2024). Enhancing Customer Segmentation in Online Transportation Services: A Comprehensive Approach Using K-Means Clustering and RFM Model. *Indonesian Interdisciplinary Journal of Sharia Economics (IIJSE)*, *7*(2), 2849-2865. https://doi.org/10.31538/iijse.v7i2.4851