Asian Journal of
Science, Technology, Engineering, and Art

e-ISSN : 3025-4507
p-ISSN : 3025-5287
Index : Harvard, Boston, Sydney University, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/AJSTEA.v2i6.3938

Data Speaks: District Clustering Map to Reveal Basic Education Problems in Samarinda City

Chiko Zet Oroh[1], Andrea Tri Rian Dani[2], Chusnul Muawanah[3], Jesselin Paskalis Sitinjak[4], Muhammad Irsadul Ibaad[5], Mislan[6], Raditya Arya Kosasih[7]
[1,2,4,5,6,7]Mulawarman University, Samarinda, Indonesia
[3]Education and Culture Department of Samarinda City, Samarinda, Indonesia
andreatririandani@fmipa.unmul.ac.id

Article Info:

Submitted:	Revised:	Accepted:	Published:
Oct 29, 2024	Nov 15, 2024	Nov 25, 2024	Dec 1, 2024

Abstract

Education is an effort to create a learning environment and educational process that encourages the development of students' potential. One of the statistical methods that can be used in the field of education is cluster analysis. Average linkage is a hierarchical clustering algorithm that groups objects based on the average distance between objects in different clusters. In this study, sub-districts were grouped based on education indicators at the primary/elementary level in Samarinda City in 2023. The results of this research obtained 2 clusters that were the most optimal because they had the highest Silhouette coefficient value. Cluster 1 contains the subdistricts of Samarinda Ulu and Sungai Kunjang, while cluster 2 contains the subdistricts of North Samarinda, Palaran, Sungai Pinang, Sambutan, Loa Janan Ilir, Samarinda Ilir, Samarinda Seberang, and Samarinda Kota. Cluster 1 has an advantage over cluster 2 in several aspects of education. Cluster 1 shows a higher average in terms of the number of students, the number of classrooms, the number of institutions, the number of certified teachers, and the number of A-accredited schools at the primary/secondary level in Kota Samarinda in 2023.

Keywords: Education; Cluster Analysis; Average Linkage; Silhouette Coefficient

INTRODUCTION

Education is an effort made to create a learning environment and educational process, so that it can proactively develop the potential of students (Margiati & Puspaningtyas, 2021). Education in practice is certainly influenced by the environment, which is a space for humans to interact (Fajri, 2021). Education can be achieved through school, community and family environments (Sadewo, 2021). The current education system in Indonesia is a national education system.(Hermanto, 2020). This system aims to equip students with academic knowledge, hone cognitive skills and foster positive attitudes among students from childhood. The national education system is also value-based, where elementary school students learn about honesty, discipline and tolerance (Munirah, 2015; Safitri, 2020). Elementary school education is the basic level for students in pursuing education (Bali et al., 2022). Elementary school education contributes to building and developing students' basic knowledge which will later be used in higher education levels (Rahayu et al., 2019). In today's digital era, the amount of data generated in the education process is increasing. Statistical methods are needed as a data analysis tool to produce more precise and effective insights. One of the statistical methods that can be used in the field of education is cluster analysis. Cluster analysis is a method for grouping a set of objects into a number of clusters based on their characteristics (Dani et al., 2020; Faujia et al., 2022). The grouping is done based on the nature of similarities between objects (Dani et al., 2019; Zen et al., 2022). Objects that are in the same group will be more similar than objects between groups (Dani et al., 2019). Cluster analysis is generally divided into two types, namely hierarchical clustering and non-hierarchical clustering (Ramadani & Salma, 2022; Rembulan et al., 2020). Hierarchical clustering is used to group objects in a structured manner based on their similarity in properties where the desired number of clusters is not yet known (Laamena & Talib, 2023). There are two methods in hierarchical grouping, namely the agglomerative and divisive methods (Novidianto et al., 2020). Hierarchical clustering can use several algorithms such as single linkage, complete linkage, and average linkage (Suhaeni et al., 2018). Hierarchical grouping can be described through a dendrogram (Facendola et al., 2023). Dendrogram is a tree diagram or graph used to group objects in

cluster analysis (Alwi & Hasrul, 2018). A dendrogram is created by creating a matrix that contains the level of similarity between grouped data (Satriatama et al., 2023; Widyawati et al., 2020). Average linkage is a hierarchical clustering algorithm that groups objects based on the average distance of an object to another object in a cluster with objects in other clusters (Paramadina et al., 2019; Reinaldi et al., 2021). Average linkage is a more stable algorithm when compared to other algorithms that only consider one pair of objects (Reinaldi et al., 2021). The average linkage algorithm tends to be less affected by outliers because it uses the average value, thus helping in more accurate cluster merging (Novidianto et al., 2020).

There are several previous studies that have studied the average linkage algorithm. Zikir et al., (2022) in his research discusses the grouping of small and medium industries in Wajo Regency using the k-means algorithm and average linkage. The results of his research using the average linkage algorithm form 2 clusters, where the first cluster has 1 member and the second cluster has 29 members. Wahyuni & Jatmiko, (2019) discusses the grouping of regencies/cities based on poverty factors in Java Island using the average linkage algorithm. The results of the study formed 2 clusters, where 16.80% of regencies/cities in Java Island are included in the first cluster, while the other 83.20% are included in the second cluster. Other research conducted by Reinaldi et al., (2021), discusses the grouping of regencies/cities based on community welfare in Java Island using the single linkage, complete linkage, and average linkage algorithms. The study found that the best algorithm used was average linkage with a Silhouette Index value of 0.61.

Based on the background that has been described, the author is interested in conducting a study entitled "*Data Speaks: District Clustering Map to Reveal Basic Education Problems in Samarinda City*". This study is expected to provide useful insights for the government as a strategic consideration in efforts to improve the quality of education, especially the District area in Samarinda City in the following years.

METHODS

The data used is a publication from the Samarinda City Education and Culture Office in 2023. The sampling technique in the study was purposive sampling, which is considering samples that have the information needed by researchers. In this study, the consideration was the availability of the latest data at the Samarinda City Education and Culture Office.

The data analysis technique used in this study was grouping areas based on Districts in Samarinda City with hierarchical grouping of the average linkage algorithm applied to the elementary school/equivalent education indicators in Samarinda City.

Software used in the data analysis process, namely Microsoft Excel and Rstudio. The analysis steps can be described as follows:

1. Conducting data exploration with spatial mapping.
2. Perform data standardization using Z-Score standardization.

$$X_{ip}^{*} = \frac{X_{ip} - \bar{X}_{p}}{S_{p}} \tag{1}$$

3. Calculates the Euclidean distance matrix.

$$d_{uv} = \sqrt{\sum_{p=1}^{j} \left(X_{up} - X_{vp} \right)^{2}} \tag{2}$$

4. Perform cluster analysis using the average linkage algorithm.

$$d_{(UV)W} = \frac{\sum_{i=1}^{n} \sum_{b=1}^{n} d_{ib}}{N_{UV} N_{W}} \tag{3}$$

5. Merge two clusters with the minimum distance.
6. Determining the optimal cluster using the Silhouette coefficient.

$$SC = maks_{k} SI(k) \tag{4}$$

7. Profiling and interpretation of cluster results.

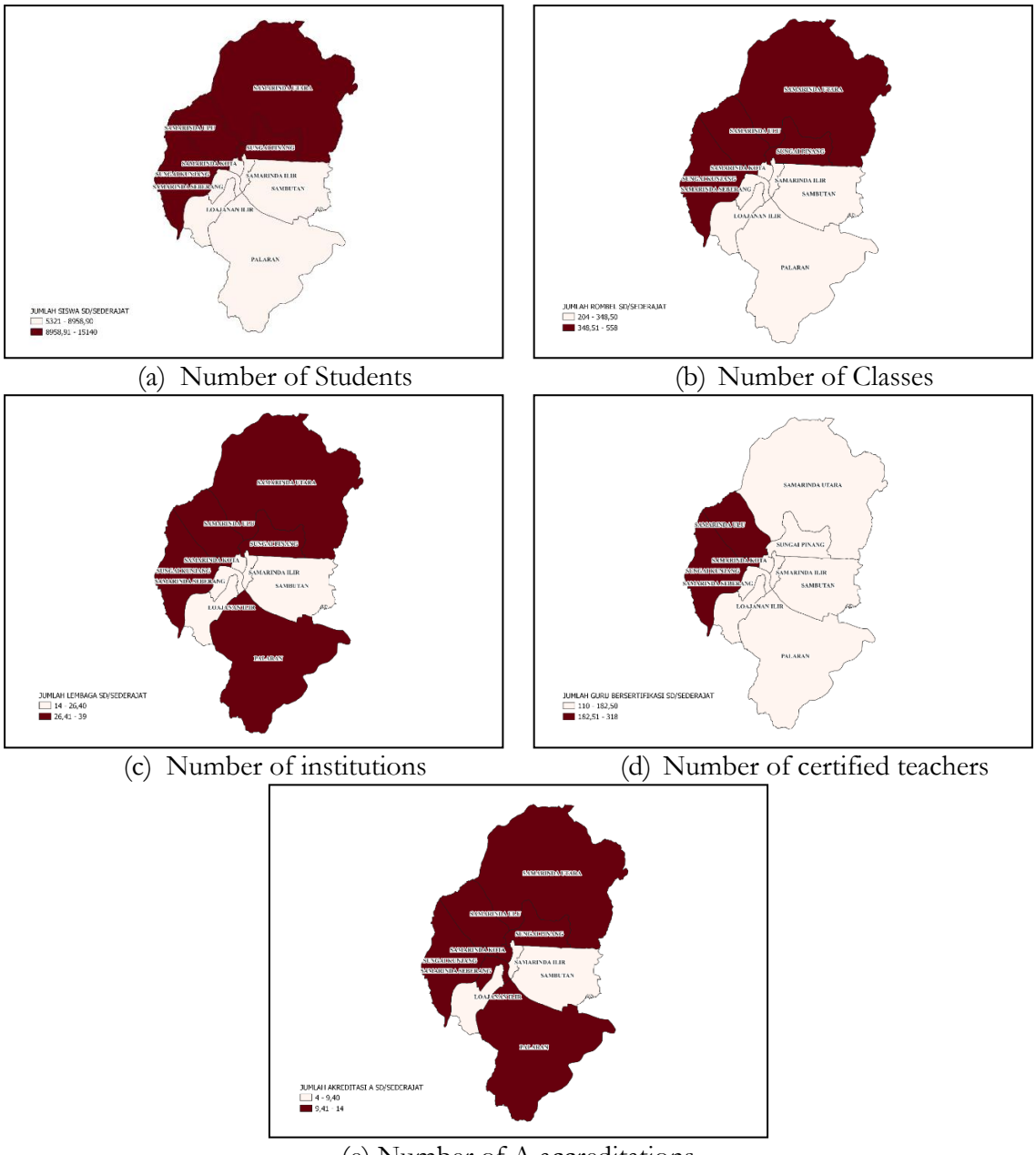
RESULTS

This study used secondary data, namely data on education indicators at the elementary school/equivalent level in the sub-district area in Samarinda City which were summarized by the Samarinda City Education and Culture Office in 2023. There are 5 variables used, namely the variable number of elementary school/equivalent students (X_1), number of study groups (rombel) for elementary school/equivalent (X_2), number of elementary/equivalent institutions (X_3), number of certified elementary school/equivalent teachers (X_4), number of A accreditation of elementary schools/equivalent (X_5).

Grouping of sub-district areas in Samarinda City using the average linkage algorithm using the five variables studied.

Spatial Mapping

Spatial mapping used to show color degradation based on its level. Spatial mapping on each research variable uses two categories, namely the low category when the value is below the average and the high category when the value is above the average. Spatial mapping is created using QGIS software.



(a) Number of Students (b) Number of Classes

(c) Number of institutions (d) Number of certified teachers

(e) Number of A accreditations

Figure 1 : Spatial mapping Elementary/equivalent education indicators in Samarinda City

Based on Figure 1(a), it can be seen that Samarinda Ulu, Samarinda Utara, Samarinda Kota, Samarinda Seberang, Sungai Kunjang, and Sungai Pinang Districts have a higher number of elementary school students than the average number of elementary school students in Samarinda City of 8,959. A large number of students can increase human resources and if managed properly will produce an educated young generation who can contribute to regional development, but this increase in the number of students can also cause the quality of education to decline if it is not balanced with an increase in school capacity and resources which can cause academic achievement to decline.

Based on Figure 1(b), it can be seen that Samarinda Ulu, North Samarinda, Samarinda City, Samarinda Seberang, Sungai Kunjang, and Sungai Pinang Districts have a higher number of elementary school/equivalent classes than the average number of elementary school/equivalent classes in Samarinda City of 349. This indicates that there has been an increase in the capacity to accommodate students in these areas. The increase in classes in an area provides an opportunity for many children to get basic education, so that it can encourage an increase in school participation rates in the area. The challenge of this increase in classes is the availability of teaching staff, facilities, and school management if not managed properly, it is possible that the quality of education in the area will be threatened.

Based on Figure 1(c), it can be seen that Sungai Kunjang, Sungai Pinang, Samarinda City, Samarinda Ulu, North Samarinda, and Palaran Districts have a higher number of elementary/equivalent schools than the average number of elementary/equivalent schools in Samarinda City of 27. The number of elementary/equivalent schools that is higher than the average in a region indicates that there are many elementary/equivalent schools operating in the region. The large number of educational institutions can create competition between schools so that it can improve the quality of education in the region. The large number of educational institutions is also often found to have disparities in facilities and adequate teaching staff, creating differences between high-quality schools and low-quality schools.

Based on Figure 1(d), it can be seen that Sungai Kunjang District, Samarinda City, Samarinda Ulu has a number of certified elementary school/equivalent teachers that is higher than the average number of certified elementary school/equivalent teachers in Samarinda City of 183. The number of certified elementary school/equivalent teachers that

is higher than the average indicates that many teachers in the area have met the competency standards set by the government. Certified teachers generally have better pedagogical and scientific competencies, so they are able to provide a more effective and high-quality learning process.

Based on Figure 1(e), it can be seen that Samarinda Seberang, Samarinda Kota, Samarinda Ulu, Samarinda Utara, Sungai Kunjang, Sungai Pinang, and Palaran Districts have a higher number of A accreditation elementary schools/equivalent than the average number of A accreditation elementary schools/equivalent throughout Samarinda City of 10. Schools with A accreditation are considered to have very good quality in various aspects, such as curriculum, management, facilities and infrastructure, and teacher competence. Areas that have many schools with A accreditation are often known as areas with superior education systems. Schools with A accreditation may experience overpopulation (excess students) because many parents want to send their children to these schools.

Data Standardization

Data standardization is carried out before grouping is carried out in order to provide more accurate grouping results. In this study, Z-Score standardization is used so that the data will be in the same value range with an average of 0 and a standard deviation of 1. Standardization of observation data for the variable number of elementary school students (X_1) using Equation (1).

$$X_{11}^{*} = \frac{13.728 - 8.958,90}{3237,05} = 1,47$$

$$X_{21}^{*} = \frac{6.515 - 8.958,90}{3237,05} = -0,75$$

$$\vdots$$

$$X_{101}^{*} = \frac{6.515 - 8.958,90}{3237,05} = 0,15$$

The standardization calculation is carried out in the same way until the variable number of elementary schools that have been accredited A (X_5). The results of data standardization can be seen in Table 1.

Table 1. Data Standardization Results with Z-Score

I	X_1^*	X_2^*	X_3^*	X_4^*	X_5^*
1	1.47	1.48	1.34	1.37	0.54
2	-0.75	-0.94	-1.32	-0.55	-0.47
3	-0.46	-0.58	-0.68	-0.30	0.20
4	-0.58	-0.28	0.49	-0.28	0.54
5	1.91	1.66	1.02	2.17	1.54
6	0.46	0.71	1.34	-0.25	0.20
7	-0.56	-0.53	-0.79	-0.18	0.20
8	-1.12	-1.14	-0.68	-1.16	-1.81
9	-0.51	-0.73	-0.89	-0.73	-1.47
10	0.15	0.37	0.17	-0.09	0.54

Average Linkage Algorithm Grouping

Grouping of Sub-districts in Samarinda City using the average linkage algorithm is done by calculating the Euclidean distance matrix using Equation (2). The Euclidean distance matrix obtained from the standardized data is as follows:

Table 2. Euclidean Distance Matrix of Samarinda City

$d_{euclidean}$	SUl	Sir	SSb	Pl	SiK	SUt	SKt	Sb	LJI	SiP
SUl	0									
Sir	4.76	0								
SSb	3.87	1.07	0							
Pl	3.28	2.20	1.26	0						
SiK	1.40	5.55	4.62	4.15	0					
SUt	2.09	3.44	2.57	1.70	3.27	0				
SKt	3.90	1.04	0.19	1.35	4.63	2.67	0			
Sb	5.44	1.66	2.35	2.94	6.49	3.86	2.38	0		
LJI	4.72	1.15	1.75	2.52	5.72	3.32	1.77	0.94	0	
SiP	2.55	2.44	1.47	1.04	3.40	1.31	1.53	3.35	2.69	0

Description: SUl: Samarinda Ulu; SIr: Samarinda Ilir; SSb: Samarinda Seberang; Pl: Palaran; SKt: Samarinda City; Sb: Welcome; LJI: Loa Janan Ilir; SiP: Sungai Pinang.

Based on the Euclidean matrix Table 2, two objects that have the minimum distance will be combined into 1 cluster such as Samarinda Seberang District with Samarinda City.

The iteration continues untilall observations are combined into 1 cluster so thata dendrogram was obtained as in Figure 2.

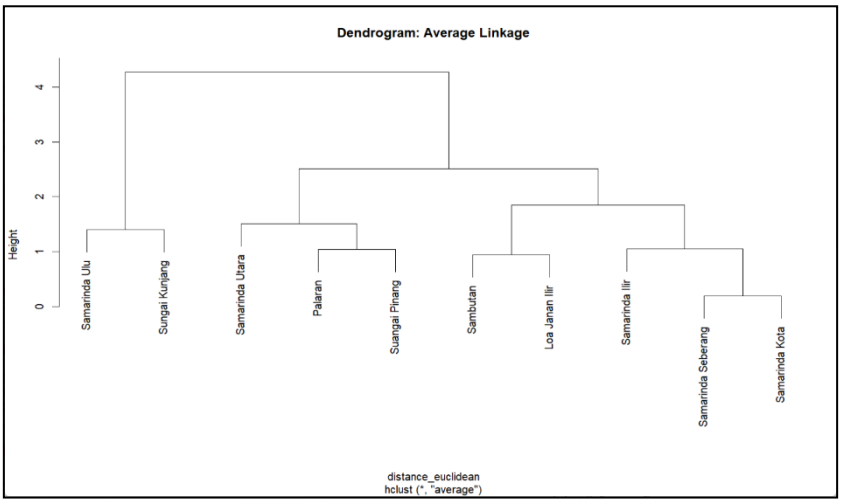
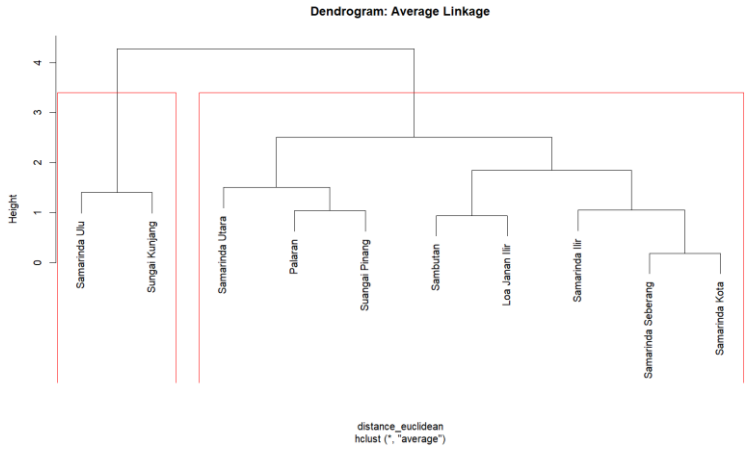
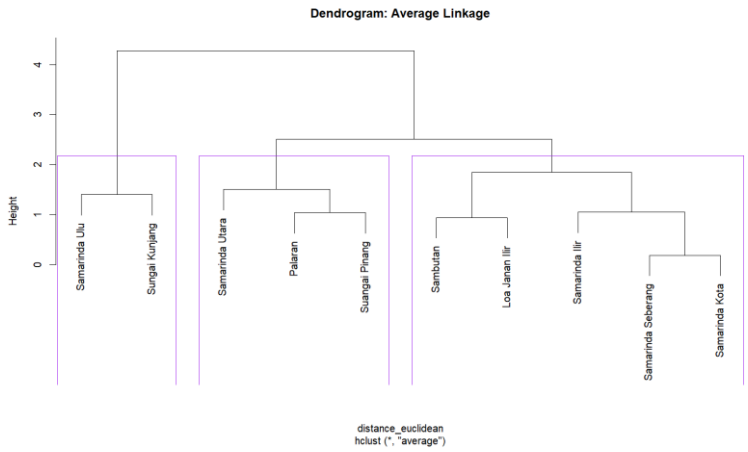
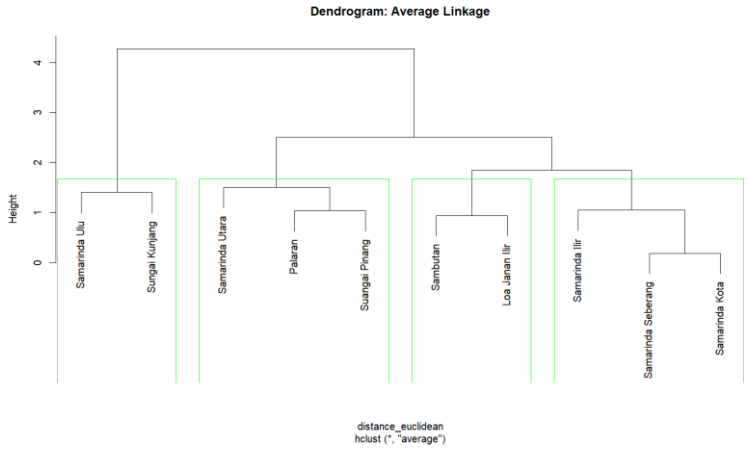
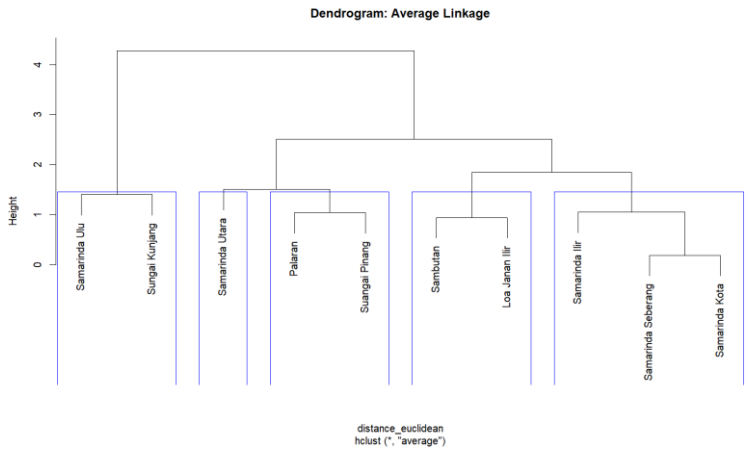


Figure 2 : Clustering results with the average linkage algorithm

Grouping based on the average linkage algorithm was tried for several numbers of clusters, namely 2, 3, 4, and 5. The results of grouping for several numbers of clusters with the average linkage algorithm can be seen in Table 3.

Table 3. Dendrogram and Interpretation of Number of Clusters

Dendrogram	Interpretation of Cluster Count
 (a) $k = 2$	Grouping with 2 clusters: - Cluster 1 contains the districts of Samarinda Ulu and Sungai Kunjang. - Cluster 2 consists of the districts of North Samarinda, Palaran, Sungai Pinang, Sambutan, Loa Janan Ilir, Samarinda Ilir, Samarinda Seberang, and Samarinda City.

Dendrogram	Interpretation of Cluster Count
 (b) $k = 3$	Grouping with 3 clusters: - Cluster 1 contains the districts of Samarinda Ulu and Sungai Kunjang. - Cluster 2 consists of the districts of North Samarinda, Palaran, and Sungai Pinang. - Cluster 3 consists of the districts of Sambutan, Loa Janan Ilir, Samarinda Ilir, Samarinda Seberang, and Samarinda City.
 (c) $k = 4$	Grouping with 4 clusters: - Cluster 1 contains the districts of Samarinda Ulu and Sungai Kunjang. - Cluster 2 consists of the districts of North Samarinda, Palaran, and Sungai Pinang. - Cluster 3 contains the districts of Sambutan and Loa Janan Ilir. - Cluster 4 consists of the districts of Samarinda Ilir, Samarinda Seberang, and Samarinda City.
 (d) $k = 5$	Grouping with 5 clusters: - Cluster 1 contains the districts of Samarinda Ulu and Sungai Kunjang. - Cluster 2 contains North Samarinda District. - Cluster 3 contains Palaran and Sungai Pinang Districts. - Cluster 4 contains the districts of Sambutan and Loa Janan Ilir. - Cluster 5 consists of the districts of Samarinda Ilir, Samarinda Seberang, and Samarinda City.

Silhouette Coefficient

The Silhouette coefficient is used to determine the optimal cluster of an algorithm used. Determining the optimal cluster with the Silhouette coefficient using Rstudio software can be seen in Figure 3.

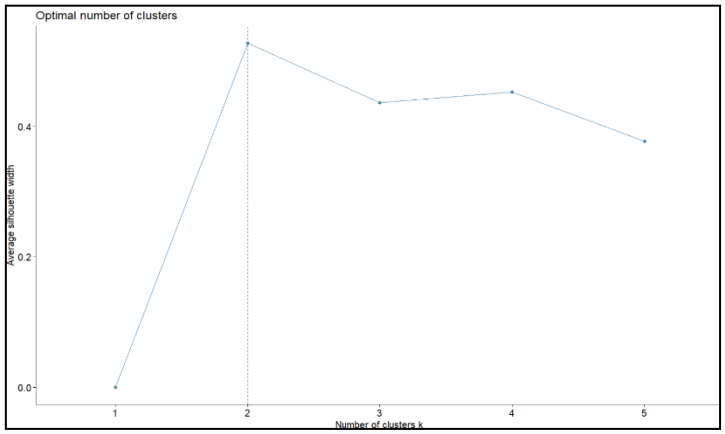


Figure 3: Determination of optimal clusters with Silhouette coefficient

Based on Figure 3 above, it can be seen that the most optimal number of clusters is $k = 2$ because it has the highest average value. Therefore, the number of clusters used is 2 clusters to group the Districts based on the elementary school/equivalent education indicators in Samarinda City in 2023.

DISCUSSION

Grouping based on the average linkage algorithm and the Silhouette coefficient value obtained the most optimal number of clusters, which is 2 clusters. The profiling of the grouping results can be seen in Table 3.

Table 4. Grouping Results with Average Linkage Algorithm $k = 2$

Cluster	Subdistrict
1	Samarinda Ulu and the Kunjang River
2	The areas covered by this program are North Samarinda, Palaran, Sungai Pinang, Sambutan, Loa Janan Ilir, Samarinda Ilir, Samarinda Seberang, and Samarinda City.

To compare the two clusters, we can see the average value of each variable used in each cluster, namely the number of students, number of classes, number of institutions,

number of certified teachers, and number of A accreditation of elementary schools/equivalent in Samarinda City.

Table 5. Cluster Profiling

Cluster	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	$\bar{X}_4$	$\bar{X}_5$
Cluster 1	14,434	547	37.50	293	12.50
Cluster 2	7,590.12	298.88	23.62	154.88	8.62

Cluster 1 has advantages over cluster 2 in several aspects of education. Cluster 1 shows a higher average in terms of the number of students, the number of classes, the number of institutions, the number of certified teachers, and the number of schools with A accreditation at the elementary school/equivalent level. The higher number of students and classes in cluster 1 indicates a more massive capacity and teaching and learning activities. The greater number of institutions can indicate that in cluster 1 there are more school choices, so that students have wider access to education. The higher number of certified teachers indicates better teaching quality, because these teachers have met professional competency standards. The greater number of A-accredited schools in cluster 1 indicates that the quality of education in schools in cluster 1 is better. This reflects better quality of education and greater capacity in organizing education in cluster 1.

CONCLUSION

Based on the results of the study of the grouping of sub-districts based on indicators of elementary school/equivalent education in Samarinda City in 2023, the following conclusions were obtained:

1. Based on the results of the grouping of sub-districts based on the elementary school/equivalent education indicator in Samarinda City in 2023 using the average linkage algorithm, the optimal number of clusters (k) using the Silhouette coefficient was obtained, namely 2 clusters.
2. The results of the grouping of Sub-districts in Samarinda City produced 2 clusters. Cluster 1 contains Samarinda Ulu and Sungai Kunjang Sub-districts, while cluster 2 contains North Samarinda, Palaran, Sungai Pinang, Sambutan, Loa Janan Ilir, Samarinda Ilir, Samarinda Seberang, and Samarinda City Sub-districts.

REFERENCES

Ahmad Zikir, Khalilah Nurfadilah, Irwan, & Adiatma. (2022). Perbandingan Metode Clustering Dengan Menggunakan Metode Average Linkage dan Metode K-Means Pada Industri Kecil dan Menengah Di Kabupaten Wajo. *Jurnal MSA (Matematika Dan Statistika Serta Aplikasinya)*, *10*(2), 57–62. https://doi.org/10.24252/msa.v10i2.31333

Alwi, W., & Hasrul, M. (2018). Analisis Klaster Untuk Pengelompokkan Kabupaten/Kota Di Provinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat. *Jurnal MSA (Matematika Dan Statistika Serta Aplikasinya)*, *6*(1), 35. https://doi.org/10.24252/msa.v6i1.4782

Bali, E. N., Bunga, B., & Kale, S. (2022). Kampus Mengajar: Upaya Transformasi Mutu Pendidikan Sekolah Dasar Di Nusa Tenggara Timur. *Jurnal Pendidikan Dasar Flobamorata*, *3*(1), 237–241. https://doi.org/10.51494/jpdf.v3i1.658

Dani, A. T. R., Wahyuningsih, S., & Rizki, N. A. (2019). Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu. *Jambura Journal of Mathematics*, *1*(2), 64–78. https://doi.org/10.34312/jjom.v1i2.2354

Dani, A. T. R., Wahyuningsih, S., & Rizki, N. A. (2020). Pengelompokkan Data Runtun Waktu menggunakan Analisis Cluster (Studi Kasus: Nilai Ekspor Komoditi Migas dan Nonmigas Provinsi Kalimantan Timur Periode Januari 2000-Desember 2016). *Jurnal EKSPONENSIAL*, *11*(1), 29–38.

Facendola, R., Ottomano Palmisano, G., De Boni, A., Acciani, C., & Roma, R. (2023). Profiling Citizens on Perception of Key Factors of Food Security: An Application of K-Means Cluster Analysis. *Sustainability (Switzerland)*, *15*(13). https://doi.org/10.3390/su15139915

Fajri, N. (2021). at-Tarbiyah al-Mustamirrah : Jurnal Pendidikan Islam Implementasi Penguatan Pendidikan Karakter di Satuan Pendidikan Sekolah Dasar. *Jurnal Pendidikan Islam*, *2*, 1–10.

Faujia, R. A., Setianingsih, E. S., & Pratiwi, H. (2022). Analisis Klaster K-Means Dan Agglomerative Nesting Pada Indikator Stunting Balita Di Indonesia. *Seminar Nasional Official Statistics*, *2022*(1), 1249–1258. https://doi.org/10.34123/semnasoffstat.v2022i1.1511

Hermanto, B. (2020). Perekayasaan sistem pendidikan nasional untuk mencerdaskan kehidupan bangsa. *Foundasia*, *11*(2), 52–59. https://doi.org/10.21831/foundasia.v11i2.26933

Laamena, N. S., & Talib, T. (2023). Penerapan Analisis Klaster Hierarki Untuk Pengelompokan Kabupaten/ Kota Di Provinsi Maluku Berdasarkan Status Pendidikan. *Science Map Journal*, *5*(1), 10–18. https://doi.org/10.30598/jmsvol5issue1pp10-18

Margiati, D. P., & Puspaningtyas, N. D. (2021). Implementasi Manajemen Pendidikan Sekolah Dasar Negri 1 Sidodadi. *Journal of Arts and Education*, *1*(1), 39–44. https://doi.org/10.33365/jae.v1i1.28

Munirah. (2015). Sistem Pendidikan di Indonesia antara Keinginan dan Realita. *Fakultas Tarbiyah Dan Keguruan UIN Alauddin Makassar*, *2*(2), 233–245.

Novidianto, R., Tri, A., Dani, R., & Hubei, P. (2020). *DI INDONESIA BERDASARKAN DATA DERET WAKTU Pada Desember Pneumonia misterius pertama kali dihubungkan*

dengan hewan kelelawar yang menular ke hewan lain sebelum menjangkiti manusia . Mulai akhir bulan Desember 2019 meningkat . Tidak sampai satu bulan , COVI. 15–24.

Paramadina, M., Sudarmin, S., & Aidid, M. K. (2019). Perbandingan Analisis Cluster Metode Average Linkage dan Metode Ward (Kasus: IPM Provinsi Sulawesi Selatan). *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, *1*(2), 22. https://doi.org/10.35580/variansiunm9357

Rahayu, P., Resnani, R., & Kustianti, S. K. (2019). Pemahaman Siswa Terhadap Bahan Ajar Muatan Lokal Bahasa Daerah Bengkulu Kelas IV SDN 04 Bengkulu (Uji Coba Bahan Ajar). *JURIDIKDAS: Jurnal Riset Pendidikan Dasar*, *2*(1), 47–53.

Ramadani, R., & Salma, A. (2022). Metode Average Linkage Dan Ward Dalam Pengelom Pokan Kesejahteraan Sumatera Barat Tahun 2021. *Journal Of Mathematics UNP*, *7*(3), 11–24.

Reinaldi, Y., Ulinnuha, N., & Hafiyusholeh, M. (2021). Comparison of Single Linkage, Complete Linkage, and Average Linkage Methods on Community Welfare Analysis in Cities and Regencies in East Java. *Jurnal Matematika, Statistika Dan Komputasi*, *18*(1), 130–140. https://doi.org/10.20956/j.v18i1.14228

Rembulan, G. D., Wijaya, T., Palullungan, D., Alfina, K. N., & Qurthuby, M. (2020). Kebijakan Pemerintah Mengenai Coronavirus Disease (COVID-19) di Setiap Provinsi di Indonesia Berdasarkan Analisis Klaster. *JIEMS (Journal of Industrial Engineering and Management Systems)*, *13*(2). https://doi.org/10.30813/jiems.v13i2.2280

Sadewo, D. S. (2021). Hubungan Perhatian Orang Tua terhadap Motivasi Belajar Anak dalam Mengerjakan Perkerjaan Rumah. *Jurnal Pendidikan Dan Pembelajaran Sekolah Dasar*, *1*(2), 59–66.

Safitri, K. (2020). Pentingnya Pendidikan Karakter Untuk Siswa Sekolah. *Jurnal Pendidikan Tambusai*, *4*(1), 264–271.

Satriatama, A. E., Wibowo, A. P., Arnold, I. G. N., Pratama, R. B., Masyhuda, T. A., Agusti, Y. A., Purwanti, E., & Werdiningsih, I. (2023). Analisis Klaster Data Pasien Diabetes untuk Identifikasi Pola dan Karakteristik Pasien. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, *5*(3), 172–182. https://doi.org/10.47233/jteksis.v5i3.828

Suhaeni, C., Kurnia, A., & Ristiyanti, R. (2018). Perbandingan Hasil Pengelompokan menggunakan Analisis Cluster Berhirarki, K-Means Cluster, dan Cluster Ensemble (Studi Kasus Data Indikator Pelayanan Kesehatan Ibu Hamil). *Jurnal Media Infotama*, *14*(1). https://doi.org/10.37676/jmi.v14i1.469

Wahyuni, S., & Jatmiko, Y. A. (2019). Pengelompokan Kabupaten/Kota di Pulau Jawa Berdasarkan Faktor-Faktor Kemiskinan dengan Pendekatan Average Linkage Hierarchical Clustering. *Jurnal Aplikasi Statistika & Komputasi Statistik*, *10*(1), 1. https://doi.org/10.34123/jurnalasks.v10i1.197

Widyawati, W., Saptomo, W. L. Y., & Utami, Y. R. W. (2020). Penerapan Agglomerative Hierarchical Clustering Untuk Segmentasi Pelanggan. *Jurnal Ilmiah SINUS*, *18*(1), 75. https://doi.org/10.30646/sinus.v18i1.448

Zen, M. A., Wahyuningsih, S., & Dani, A. T. R. (2022). Aplikasi Pendekatan Agglomerative Hierarchical Time Series Clustering untuk Peramalan Data Harga Minyak Goreng di Indonesia. *Seminar Nasional Official Statistics*, *2022*(1), 293–302. https://doi.org/10.34123/semnasoffstat.v2022i1.1394