Jurnal Varian
Vol. 9, No. 1, February 2026, pp 67–76
e-ISSN: 2581-2017
DOI: https://doi.org/10.30812/varian.v9i1.6035

Principal Component Analysis and Agglomerative Hierarchical Clustering for Assessing the Condition of MSMEs Assisted by the Department of Cooperatives and MSMEs

Ardiana Fatma Dewi[1], **Kurnia Ahadiyah**[2]
[1]Universitas Islam Negeri Sayyid Ali Rahmatullah Tulungagung, Tulungagung, Indonesia
[2]Universitas Islam Negeri Syekh Wasil Kediri, Kediri, Indonesia

Article Info

Article history:

Received : 12-17-2025
Revised : 03-02-2026
Accepted : 03-16-2026

Keywords:

Clustering;
Economy;
Principal Component Analysis;
MSME.

ABSTRACT

Micro, Small, and Medium Enterprises (MSMEs) have an important role in the growth of the Indonesian economy. To achieve these hopes certainly requires a strategy. One way is to formulate policies based on information adapted to local conditions. One of the right ways to conduct this research is through data mining. There are techniques in data mining and one of the techniques that can be used is clustering with the Agglomerative Hierarchical Clustering Algorithm with Principal Component Analysis (PCA). Cluster analysis aims to group objects based on their characteristics. This research aims to determine the appropriate distribution strategy for business capital assistance based on the characteristics of MSMEs assisted by the Kediri City Cooperatives and MSMEs Department. In grouping MSMEs assisted by the Department of Cooperatives and MSMEs of Kediri City based on several indicators measured by business capital, turnover, profits, human resources, marketing methods, government capital assistance, type of business, and place of business, it was found that the optimal algorithm used was complete linkage. With a cophenetic correlation value of 0.733. Based on strong internal cluster validation, as indicated by silhouette values, the characteristics of MSME actors suggest 3 representative clusters. An interesting finding is that the third cluster has not had access to government assistance programs. Based on the results of this research, it can be concluded that government capital assistance is not evenly distributed and is not optimal for increasing MSME competitiveness.

Corresponding Author:

Ardiana Fatma Dewi,
Department of Mathematics Education,
Universitas Islam Negeri Sayyid Ali Rahmatullah Tulungagung,
Email: ardianafatmad@iainkediri.ac.idd

How to Cite:
Dewi, A. F., & Ahadiyah, K. (2026). Principal Component Analysis and Agglomerative Hierarchical Clustering for Assessing the Condition of MSMEs Assisted by the Department of Cooperatives and MSMEs. *Jurnal Varian*, *9*(1), 67–76.

A. INTRODUCTION

Micro, Small, and Medium Enterprises (MSMEs) play a central role in Indonesia's economic development. In addition to being the backbone of national economic growth, MSME also contributes greatly to job creation and play a role in achieving sustainable development goals (SDGs) (Putra & Hartomo, 2021; Suci, 2017). The Indonesian government has paid serious attention to empowering this sector as a strategic step to improve people's welfare (Kaswinata et al., 2023).

The rapid development of information technology opens up new opportunities in analyzing and managing MSME data more efficiently and accurately (Prabowo et al., 2023). One technology that has great potential in this regard is data mining. This technology allows the extraction of useful information from large amounts of data, which can ultimately be used as a basis for strategic decision-making (Suntoro, 2019). In the context of MSMEs, data mining can be applied to solve various business problems, such as determining marketing strategies and allocating business capital assistance (Astuti et al., 2019).

Of the various techniques in data mining, the clustering method, or data grouping, is very relevant (Suntoro, 2019). One effective algorithm for this purpose is Agglomerative Hierarchical Clustering (AHC). AHC is a grouping method that combines data based on similar characteristics to form a cluster hierarchy (Pratikto & Damastuti, 2021; Roux, 2015; Salsabila & Hendrawan, 2021). This approach can be used to identify MSME segmentation based on important attributes such as business capital, sales results, number of workers, profits, and marketing strategies (Utari & Dewi, 2014).

The implementation of this method is very important, especially for local governments such as Kediri City, which has recorded a significant increase in the number of MSMEs. By the end of 2023, there were around 5,400 MSMEs fostered by the city government spread across various sectors, including culinary, agribusiness, fashion, and crafts (Karem et al., 2024). Even so, there are significant challenges related to grouping and assessing the effectiveness of capital assistance (Huriah & Dienwati Nuris, 2023). The assistance that has been provided evenly does not take into account the specific characteristics of each MSME, so it is feared that it will be less than optimal in supporting business development (Jelita et al., 2023; Wanto et al., 2020). However, previous studies have primarily focused on clustering techniques without integrating dimensionality reduction methods such as Principal Component Analysis (PCA), and have rarely linked clustering results to the evaluation of government assistance effectiveness.

This study focuses on two main things, namely how the general condition of MSME in Kediri City is seen from aspects such as business capital, sales results, human resources, profits, and marketing, and how the right classification of MSME can be used to formulate a strategy for distributing business capital assistance from the government more effectively and in accordance with the needs of each business actor. Therefore, this study aims to group MSMEs based on key attributes to formulate a more targeted capital assistance distribution strategy. With this approach, the government aims to allocate resources more effectively and efficiently, thereby encouraging the growth and sustainability of the MSME sector. This aligns with the vision of equitable and sustainable economic development.

B. RESEARCH METHOD

This study uses primary data obtained directly from MSME actors fostered by the Cooperatives and MSME Service of Kediri City. Data collection was conducted by distributing questionnaires via Google Forms. The instrument comprised structured questions with categorical response options, allowing respondents to select predefined answer categories. The collected data primarily consisted of categorical variables, along with respondents' demographic and business identity information to support descriptive and inferential analysis. This was done to reach MSME actors across various locations in Kediri City more efficiently, as the MSME population in Kediri City is around 9,000 units. A simple random sampling technique is used to ensure that each element of the population has an equal chance of being selected into the sample. To determine the number of representative samples, the Issac and Michael formula is used with a 10% error rate. The formula used can be seen in Equation (1).

$$n = \frac{\lambda^2 NPQ}{d^2 (N-1) + \lambda^2 PQ} \tag{1}$$

n denotes the number of samples or respondents. The term λ^2 represents the Chi-Square table value (with degrees of freedom equal to 1 and an error rate of 10%). Furthermore, N refers to the total population size. The parameter P indicates the true probability, set at 0.5, while Q represents the false probability, also set at 0.5. Lastly, d denotes the difference between the sample mean and the population mean, with a value of 0.05.

Based on the formula, the minimum sample size required is 266. In its implementation, the data collected from 274 respondents were used in the analysis of this study. This study uses eight main variables to classify MSMEs, namely: business capital, sales results, profits, number of human resources (HR), marketing strategy, capital assistance status, type of business, and place of business. Each variable is categorized, such as capital and income in a certain nominal range, number of workers, marketing methods (offline, online, or mixed), status of receiving assistance (already or not), type of business (production or sales), and ownership of the place of business (privately owned or rented) (Zahrotun et al., 2023).

This study uses a quantitative approach with structured data analysis stages to produce an accurate classification of MSMEs. The initial stage begins with the preparation of a research instrument in the form of a questionnaire, which is compiled based on predetermined variables and tested through literature studies. Before being distributed to the respondents, the questionnaire was tested for validity and reliability. Validity testing was conducted to ensure that each item accurately measured the intended construct. Item validity was examined using correlation analysis between item scores and total construct scores. Items with correlation coefficients exceeding the acceptable threshold were considered valid. Furthermore, reliability testing was conducted to assess the instrument's internal consistency using Cronbach's alpha. A construct was considered reliable if the Cronbach's Alpha value exceeded

the minimum acceptable level of 0.70.

Furthermore, a descriptive statistical analysis is conducted to provide a general overview of MSME characteristics in Kediri City. For the research variables measured as categorical data, descriptive statistics included the mean, minimum, and maximum. The mean indicated the central tendency of respondents' responses, while the standard deviation reflected the data's variability or dispersion. After that, a Pearson correlation test is carried out to determine the relationship between variables, and if there is a significant correlation, it is continued with Principal Component Analysis (PCA) to reduce the number of variables without eliminating important information (Dewi & Ahadiyah, 2022; Mariana, 2013; Saraçli et al., 2013). Principal Component Analysis (PCA) was employed in this study to reduce the dataset's dimensionality and identify the underlying structure among the observed variables. PCA is a multivariate statistical technique that transforms a set of correlated variables into a smaller number of uncorrelated components called principal components. These components are linear combinations of the original variables, extracted so that the first component explains the maximum possible variance, followed by subsequent components. The next stage is calculating the distance between objects using the Euclidean formula to determine the level of data similarity between MSMEs. Equation (2) is the Euclidean distance formula used in this study (Maulidia & Wulandari, 2022).

$$d\left(x,y\right)=\sqrt{\sum_{i=1}^{p}\left(x_{ik}-y_{jk}\right)^{2}} \tag{2}$$

$d\left(x,y\right)$ represents the Euclidean distance between observation i and observation j. The term x_{ik} denotes the value of the k-th variable for observation i, while y_{jk} indicates the value of the k-th variable for observation j. Furthermore, p refers to the total number of variables or dimensions.

Agglomerative Hierarchical Clustering (AHC) is a hierarchical clustering technique that groups observations based on their similarity. The method follows a bottom-up approach, where each observation initially forms its own cluster. At each step, the two closest clusters are merged using a defined distance measure, such as the Euclidean distance, until all observations are combined into a single cluster. The clustering process is typically illustrated using a dendrogram, which visually represents the hierarchy of cluster formation and helps determine the optimal number of clusters. In this study, AHC was applied to identify homogeneous groups of respondents based on the selected variables, thereby enabling a clearer understanding of patterns within the dataset. Data grouping was done using the Agglomerative Hierarchical Clustering (AHC) method, with linkage approaches such as single, complete, average, and ward linkage (Mu'afa & Ulinnuha, 2019; Murtagh & Contreras, 2012). In this research, a complete linkage was used. The best approach was selected based on the highest cophenetic coefficient value. The formula for calculating the cophenetic correlation coefficient is displayed in Equation (3) (Saraçli et al., 2013).

$$r_{Coph}=\frac{\sum_{i=1}^{n}\left(d_{ik}-\bar{d}\right)\left(d_{c_{ik}}-\bar{d}_{c}\right)}{\sqrt{\sum_{i=1}^{n}\left(d_{ik}-\bar{d}\right)^{2}\left(d_{c_{ik}}-\bar{d}_{c}\right)^{2}}} \tag{3}$$

r_{Coph} denotes the cophenetic correlation coefficient. The term d_{ik} represents the Euclidean distance between objects i and k, while $\bar{d}$ indicates the average of d_{ik}. Furthermore, $d_{c_{ik}}$ refers to the cophenetic distance between objects i and k, and $\bar{d}c$ denotes the average of dc_{ik}.

To determine the optimal number of clusters, the Silhouette Coefficient value is used so that the resulting grouping structure is more representative (Mustika et al., 2021; Ningsih et al., 2016). The Silhouette Coefficient evaluates clustering quality by measuring how well each observation fits within its assigned cluster relative to other clusters. The coefficient ranges from -1 to 1, where values close to 1 indicate that observations are well matched to their own cluster and poorly matched to neighboring clusters. Values around 0 suggest overlapping clusters, while negative values indicate potential misclassification. Finally, an interpretation of the grouping results is provided, offering relevant strategic recommendations to the government on capital assistance and promotional strategies tailored to the characteristics of each MSME group.

C. RESULT AND DISCUSSION

1. Descriptive Analysis

The results of the descriptive analysis show that MSMEs fostered by the Kediri City Cooperative and MSME Service are generally micro-scale businesses with relatively limited initial capital, ranging from 5 to 10 million rupiah. However, these businesses can generate an average monthly turnover of 2 to 3 million rupiah. The majority of these MSMEs are managed by

less than two people, so they are very personal. Interestingly, even though the capital and scale of the business are limited, the profits are quite encouraging, around 2 to 3 million rupiah per month. This indicates that business actors are highly adaptable and innovative in managing their businesses. In terms of marketing, MSMEs in Kediri City have begun to utilize digital technology. However, most business actors have not received government capital assistance. Government support in the form of wider access to capital is still very much needed to encourage more significant business growth. This finding is in line with previous studies, which state that MSMEs often face limitations in access to capital and rely on internal capabilities to sustain their business (Utari & Dewi, 2014; Wanto et al., 2020).

2. Agglomerative Hierarchical Clustering (AHC)

a. Correlation Test

A correlation test is one of the statistical tools used to measure the strength and direction of the relationship between two numeric variables (Bouguettaya et al., 2015). In other words, this test helps us understand how closely two variables are related. The initial step before conducting cluster analysis is to run a multicollinearity test, as the variables used in this study are quite numerous. The test for multicollinearity between variables involves examining the correlations among all research variables. The results of the correlation test are shown in the Pearson correlation value. The results of the Matrix Correlation Test can be seen in Table 1.

Table 1. Results of the Matrix Correlation Test of Research Variables

Variable	Turnover	HR	Profit	Marketing	Capital Assistance	Business	Place
Capital	0.000	0.000	0.000	0.059	0.466	0.026	0.000
Turnover		0.000	0.000	0.450	0.416	0.876	0.000
HR			0.000	0.000	0.911	0.726	0.001
Profit				0.399	0.313	0.847	0.000
Marketing					0.291	0.043	0.665
Capital Assistance						0.000	0.018
Business							0.716

b. Principal Component Analysis

The Kaiser-Meyer-Olkin (KMO) Test and the Bartlett Test are two very important statistical tests in conducting principal component analysis (PCA). Both of these tests are used to determine whether the data we have is suitable for factor analysis or principal components analysis. Before conducting Principal Component Analysis, a test needs to be conducted to determine the adequacy of the data, namely the Kaiser Meyer Olkin (KMO) value, which aims to determine whether all the data taken is sufficient to be factored, and the Bartlett's sphericity assumption test is carried out to determine whether there is a relationship between variables.

Table 2. Kaiser-Mayer Olkin (KMO) Test

Test	Value
Kaiser–Meyer–Olkin Measure of Sampling Adequacy	0.688
Bartlett's Test of Sphericity – Approx. Chi-Square	412.537
Bartlett's Test of Sphericity – df	28
Bartlett's Test of Sphericity – Sig.	.000

Based on Table 2, the KMO value is 0.688. If KMO$>$ 0.5, then it can be said that the analyzed data has met the assumption of data sufficiency (Hill, 2012). The next test is Bartlett's test, which tests the null hypothesis that the population correlation matrix is the identity matrix, i.e., there is no correlation between variables. The following is the Bartlett test on MSME data.

Table 3. Bartlett Spherecity Test

	Chi-Square	p-value
Bartlett Spherecity	514.89	2.2×10^{-16}

Based on Table 3, the Chi-Square value of 514.89 with a significance level of 5% yields a p-value of 2.2×10^{-16}. If the p-value is less than 5%, then it can be concluded that there is a correlation between variables in the data of MSMEs fostered

by the Cooperatives and MSME Service in Kediri City. After examining the assumptions of principal component analysis, a factor analysis is performed to determine the new factors. The following is a reference used to construct a factor analysis for the table below.

Table 4. Eigen Value

Value	Component					
	1	2	3	4	...	8
Eigen Value	1.631	1.137	1.072	0.937		0.447
% of Variances	0.332	0.161	0.144	0.109		0.025
Comulative %	0.332	0.494	0.640	0.747		1

Based on Table 4, it can be seen that when the eigenvalue is greater than 1, the component has many factors that will later form. namely the 3rd component. This means that in this study. from the ten variables used. It will be simplified into 3 factors. with a total variance of 64%. The new factors formed are free from correlation and will then be subjected to cluster analysis (Santoso, 2018). The results of the principal component analysis are principal component scores, calculated using the principal component equation. and will then be used as input to replace the original variables. The following are the results of the correlation after the Principal component analysis is shown in the table below.

Table 5. Principal Component Score Correlation Test

Variable	Turnover	HR	Profit	Marketing	Capital Assistance	Business	Place
Capital	0.390	0.090	0.451	0.335	0.560	0.574	0.629
Turnover		0.642	0.913	0.971	0.954	0.763	0.760
HR			0.920	0.579	0.972	0.730	0.692
Profit				0.924	0.930	0.844	0.970
Marketing					0.626	0.186	0.778
Capital Assistance						0.969	0.601
Business							0.774

Based on Table 5, it can be seen that with the principal component analysis. The data will be transformed so that the variables that initially show high correlation become less correlated. This is in accordance with the assumption in cluster analysis. that the variables are mutually independent.

c. Cluster Analysis

Cluster analysis in this study is used to group MSME actors fostered by the Cooperatives and MSME Service of Kediri City based on indicators derived from factor analysis. Grouping is based on Euclidean distance, with cluster formation using the Single Linkage algorithm, which produces the highest chopenetic correlation. Complete Linkage. Average Linkage, Ward Linkage, Centroid Linkage, Median Linkage (Asiska et al., 2019). The results of several cluster algorithms that have been mentioned can be seen in Table 6.

Table 6. Chopenetic Correlation Coefficient Value

Method	Chopenetic Correlation
Single Linkage	0.260
Complete Linkage	0.733
Average Linkage	0.623
Ward Linkage	0.509
Centroid Linkage	0.454
Median Linkage	0.381

Based on Table 6, the highest chopenetic correlation value is 0.733 for the Complete Linkage algorithm, so it is used in this study. This result is consistent with previous studies that show complete linkage often produces better cluster structures in hierarchical clustering (Asiska et al., 2019). After the optimal algorithm in cluster analysis is known. The next step is to determine the optimal number of clusters based on internal validation. The following is the result of internal cluster validation, as measured by the silhouette coefficient.

Table 7. Silhouette Coefficient Value

Value	Number of Cluster			
	3	4	5	6
Silhouette	0.2845	0.2489	0.2198	0.1989

The silhouette coefficient value ranges from -1 to 1. A value close to 1 indicates that the object is very well-suited to its cluster and very dissimilar to other clusters. This means that the grouping is very good. In Table 7, it can be seen that the grouping is of 3 to 6 clusters. The largest silhouette value that approaches 1 is in 3 clusters. So that in the grouping of MSMEs fostered by the Cooperatives and MSME Service of Kediri City, based on several indicators. They are grouped into 3 clusters.

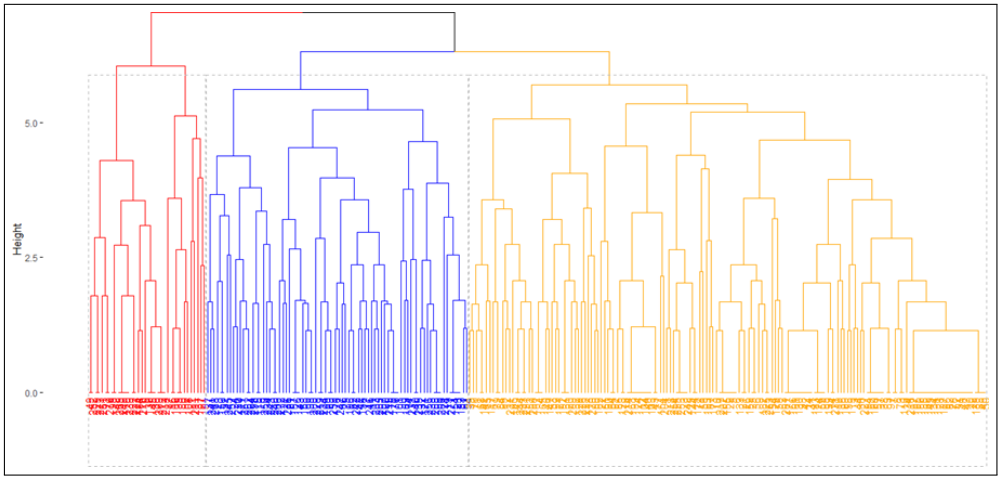

Figure 1. Dendogram Analysis of Complete Linkage Agglomerative Clustering

Based on Figure 1, the grouping of MSMEs fostered by the Cooperatives and MSME Service of Kediri City consists of 3 clusters. In the dendogam formed. The members of each cluster can also be seen. where the cluster marked in red has 29 members. The cluster marked in blue has 106 members, and the cluster marked in yellow has 139 members. The following is a profile table of the members in each cluster, as shown in the table below.

Table 8. Profile of Each Cluster

Variable	Red Cluster	Blue Cluster	Yellow Cluster
Capital	5-10 million	5-10 million	$<$ 5 million
Turnover	$>$ 3 million	$>$ 3 million	2-3 million
HR	$>$ 2 million	$>$ 2 million	1-2 million
Profit	$>$ 5 people	2-5 people	$<$ 2 people
Marketing	Digital	Mixed	Mixed
Capital Assistance	Already Received	Already Received	Not Yet Received
Business	Production	Sales	Production
Place	Rent	Private Ownership	Private Ownership

Based on Table 8, the results of the grouping of MSMEs fostered by the Kediri City Cooperative and MSME Service show some information. Among them are in the first cluster with a red line, with a total of 29 MSME actors with an average capital of 5-10 million rupiah. a monthly turnover of more than 3 million rupiah. a monthly profit of more than 2 million rupiah. A number of employees exceeding 5, with a digital marketing method, have received capital assistance from the government. type of production business. and a place of business rented.

In the second cluster, in blue, are 106 MSME actors with an average capital of 5-10 million rupiah. a monthly turnover of more than 3 million rupiah. a monthly profit of more than 2 million rupiah. a number of employees ranging from 2 to 5. With a mixed marketing approach, they have received government capital assistance. type of production business. and a privately owned place of business.

In the cluster with the most members. namely, the yellow cluster, comprising 139 MSME actors with an average capital of less than 5 million rupiah. a monthly turnover of 2-3 million rupiah. a profit of 1-2 million rupiah. The number of employees is less than 2. The marketing approach is a mix of traditional and digital. has never received government capital assistance for a production business with privately owned business premises. This finding is consistent with previous studies, which indicate that MSMEs with lower economic capacity tend to have limited access to external funding and government support (Astuti et al., 2019; Wanto et al., 2020).

D. CONCLUSION AND SUGGESTION

In the grouping of MSMEs fostered by the Kediri City Cooperative and MSME Service, based on several indicators measured from business capital. turnover. profit. HR. marketing methods. government capital assistance. type of business. and place of business using the Agglomerative Hierarchical Clustering method. It was found that the optimal algorithm used was the complete linkage algorithm. with a cophenetic correlation value obtained of 0.733. Based on cluster validation. Both internal through silhouette values based on the characteristics of MSME actors. The number of representative clusters is 3. In each group, the number of members differs. The first cluster consists of 29 MSME actors. The second cluster of 106 MSME actors. and the third cluster, 139 MSME actors.

The analysis shows a significant gap between the third cluster and the other two clusters. With a total of 139 members of micro-scale MSME actors. This cluster has an average capital. turnover. and profit that is much lower than the other two clusters. An interesting finding is that the third cluster has not yet received access to government assistance programs. Based on the results of this study. It can be concluded that government capital assistance has not been fully distributed and has not been optimal in achieving the goal of increasing MSME competitiveness. It could be that one of the main causes of this problem is the lack of an effective mechanism for identifying and targeting MSMEs that really need assistance.

ACKNOWLEDGEMENT

We would like to express our deepest appreciation to IAIN Kediri for the academic support and resources provided throughout this study. Our gratitude also extends to the Ministry of Religious Affairs for facilitating the research process, and to the Department of Cooperative and MSME of Kediri City for their valuable assistance and cooperation in providing essential data and field access.

DECLARATIONS

AUTHOR CONTRIBUTION

The contributions to this work are distributed as follows: Ardiana Fatma Dewi handled the conceptualization, writing of the original draft, editing, and visualization. Meanwhile, Kurnia Ahadiyah was responsible for the writing (review and editing), formal analysis, and the development of the study's methodology.

FUNDING STATEMENT

This research was funded by the 2024 Litapdimas research grant. The funding body had no role in the study's design or collection. analysis. or interpretation of data; or in the writing of the manuscript. Additional academic support and resources were provided by IAIN Kediri. Field access and essential data were facilitated with the cooperation of the Department of Cooperative and MSME of Kediri City. under the coordination of the Ministry of Religious Affairs.

COMPETING INTEREST

The authors declare that there are no conflicts of interest involved in this study.

REFERENCES

Asiska, N., Satyahadewi, N., & Perdana, H. (2019). Pencarian Cluster Optimum Pada Single Linkage, Complete Linkage dan Average Linkage. *Bimaster : Buletin Ilmiah Matematika, Statistika dan Terapannya*, *8*(3), 393–398. https://doi.org/10.26418/bbimst.v8i3.33173

Astuti, D., Iskandar, A. R., & Febrianti, A. (2019). Penentuan Strategi Promosi Usaha Mikro Kecil dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering. *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, *1*(2), 60–72. https://doi.org/10.20895/inista.v1i2.71

Bouguettaya, A., Yu, Q., Liu, X., Zhou, X., & Song, A. (2015). Efficient agglomerative hierarchical clustering. *Expert Systems with Applications*, *42*(5), 2785–2797. https://doi.org/10.1016/j.eswa.2014.09.054

Dewi, A. F., & Ahadiyah, K. (2022). Agglomerative Hierarchy Clustering Pada Penentuan Kelompok Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Pendidikan. *Zeta - Math Journal*, *7*(2), 57–63. https://doi.org/10.31102/zeta.2022.7.2.57-63

Hill, B. D. (2012). *The Sequential Kaiser-Meyer-Olkin Procedure as an Alternative for Determining the Number of Factors in Common-Factor Analysis: A Monte Carlo Simulation*. Proquest, Umi Dissertation Publishing.

Huriah, D. A., & Dienwati Nuris, N. (2023). Klasifikasi Penerima Bantuan Sosial Umkm Menggunakan Algoritma Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 360–365. https://doi.org/10.36040/jati.v7i1.6300

Jelita, T., Buaton, R., & Simanjuntak, M. (2023). Pengelompokan Bidang Usaha Terhadap Bantuan Produktif Usaha Mikro (BPUM) Berdasarkan Wilayah Deli Serdang Menggunakan Metode Clustering K-Means (Studi Kasus: Dinas Koperasi Dan UMKM Kabupaten Deli Serdang). *Explorer*, *3*(2), 50–57. https://doi.org/10.47065/explorer.v3i2.783

Karem, N. A., Yuliani, Y., & Mutafarrida, B. (2024). Strategi Pemberdayaan Usaha Mikro, Kecil dan Menengah (UMKM) di Kota Kediri. *Jurnal Ilmiah Nusantara*, *1*(4), 149–162. https://doi.org/10.61722/jinu.v1i4.1683

Kaswinata, K., Harahap, I., Nawawi, Z. M., & Syahputra, A. (2023). Signifikansi Perananan UMKM dalam Pembangunan Ekonomi di Kota Medan Dalam Perspektif Syariah. *Jurnal Tabarru': Islamic Banking and Finance*, *6*(2), 718–728. https://doi.org/10.25299/jtb.2023.vol6(2).15302

Mariana, M. (2013). Analisis Komponen Utama. *Matematika dan Pembelajaran*, *1*(2), 189–204. https://doi.org/10.33477/mp.v1i2.304

Maulidia, N., & Wulandari, S. P. (2022). Analisis Cluster dan Korespondensi terhadap Indikator Pertumbuhan Penduduk Kota Surabaya Tahun 2020. *Jurnal Sains dan Seni ITS*, *11*(1), D43–D49. https://doi.org/10.12962/j23373520.v11i1.62843

Mu'afa, S. F., & Ulinnuha, N. (2019). Perbandingan Metode Single Linkage, Complete Linkage Dan Average Linkage dalam Pengelompokan Kecamatan Berdasarkan Variabel Jenis Ternak Kabupaten Sidoarjo. *Inform : Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, *4*(2). https://doi.org/10.25139/inform.v4i2.1696

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: An overview. *WIREs Data Mining and Knowledge Discovery*, *2*(1), 86–97. https://doi.org/10.1002/widm.53

Mustika, M., Ardilla, Y., Manuhutu, A., Ahmad, N., Hasbi, I., Guntoro, G., Manuhutu, M. A., Ridwan, M., Hozairi, H., Wardhani, A. K., Alim, S., Romli, I., Religia, Y., Octafian, D. T., Sufandi, U. U., & Ernawati, I. (2021). *Data Mining dan Aplikasinya*. Penerbit Widina.

Ningsih, S., Wahyuningsih, S., & Nasution, Y. N. (2016). Perbandingan Kinerja Metode Complete Linkage dan Average Linkage dalam Menentukan Hasil Analisis Cluster (Studi Kasus: Produksi Palawija Provinsi Kalimantan Timur 2014/2015). *Prosiding Seminar Sains Dan Teknologi FMIPA Unmul*, *1*(1), 46–50.

Prabowo, O. H., Merthayasa, A., & Saebah, N. (2023). Pemanfaatan teknologi informasi dan manajemen perubahan pada kegiatan bisnis di era globalisasi. *Syntax Idea*, *5*(7), 883–892. https://doi.org/10.46799/syntax-idea.v5i7.2522

Pratikto, R. O., & Damastuti, N. (2021). Klasterisasi Menggunakan Agglomerative Hierarchical Clustering untuk Memodelkan Wilayah Banjir. *JOINTECS (Journal of Information Technology and Computer Science)*, *6*(1), 13. https://doi.org/10.31328/jointecs.v6i1.1473

Putra, A. C., & Hartomo, K. D. (2021). Optimalisasi Penyaluran Bantuan Pemerintah Untuk UMKM Menggunakan Metode Fuzzy C-Means. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *5*(3), 474–482. https://doi.org/10.29207/resti.v5i3.2980

Roux, M. (2015, September 5). *A comparative study of divisive hierarchical clustering algorithms*. arXiv: 1506.08977 [cs]. https://doi.org/10.48550/arXiv.1506.08977

Salsabila, D., & Hendrawan, M. Y. (2021). Analisis kondisi pemberdayaan gender di indonesia tahun 2020 dengan agglomerative hierarchical clustering dan biplot. *Seminar Nasional Official Statistics*, *2021*(1), 204–213. https://doi.org/10.34123/semnasoffstat.v2021i1.839

Santoso, S. (2018, July 23). *Mahir statistik multivariat dengan spss*. Elex Media Komputindo.

Saraçli, S., Doğan, N., & Doğan, İ. (2013). Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications*, *2013*(1), 203. https://doi.org/10.1186/1029-242X-2013-203

Suci, Y. R. (2017). Perkembangan Umkm (usaha Mikro Kecil Dan Menengah) Di Indonesia. *Jurnal Ilmiah Cano Ekonomos*, *6*(1), 51–58. https://doi.org/10.30606/cano.v6i1.627

Suntoro, J. (2019, May 20). *Data mining: Algoritma dan implementasi dengan pemrograman php*. Elex Media Komputindo.

Utari, T., & Dewi, N. P. M. (2014). Pengaruh Modal, Tingkat Pendidikan dan Teknologi Terhadap Pendapatan Usaha Mikro Kecil dan Menengah (UMKM) di Kawasan Imam Bonjol Denpasar Barat. *E-Jurnal Ekonomi Pembangunan*, *3*(12), 576–585. https://ojs.unud.ac.id/index.php/eep/article/view/9916

Wanto, A., Siregar, M. N. H., Windarto, A. P., Windarto, A. P., Ginantra, N. L. W. S. R., Napitupulu, D., Negara, E. S., Dewi, M. R. L. S. V., & Prianto, C. (2020). *Data Mining: Algoritma dan Implementasi*. Yayasan Kita Menulis.

Zahrotun, L., Nugroho, S. H., Linarti, U., & Jones, A. H. S. (2023). Analisis Persebaran UMKM Bidang Jasa Menggunakan Metode AHC Complete Linkage. *Kesatria J. Penerapan Sist. Inf.(Komputer dan Manajemen)*, *4*(2), 255–265.

[This page is intentionally left blank.]

Vol. 9, No. 1, February 2026, pp 67–76
DOI: https://doi.org/10.30812/varian.v9i1.6035