BAREKENG: Journal of Mathematics and Its Applications
September 2025 Volume 19 Issue 3 Page 1779-1790
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss3pp1779-1790

IMPLEMENTATION OF THE DBSCAN ALGORITHM FOR CLUSTERING STUNTING PREVALENCE TYPOLOGY IN WEST JAVA, CENTRAL JAVA, AND EAST JAVA REGIONS

Bagus Sumargo[1*], Kadir[2], Dena Safariza[3], Munawar Asikin[4], Dania Siregar[5], Nilam Novita Sari[6], Danu Umbara[7], Rizky Hilmianto[8], Robert Kurniawan[9], Irman Firmansyah[10]

[1,3,5,6,7,8] *Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta*
Jln. Rawamangun Muka Raya No. 11, Jakarta, 13220, Indonesia

[2] *Economist, BPS Statistics-Indonesia*
Jln. Dr. Soetomo No 6-8, Jakarta, 10710, Indonesia

[4] *Department of Economic, Universitas Al Azhar*
Jln. Sisingamangaraja, Jakarta, 12110, Indonesia

[9] *Department of Statistical Computing, Politeknik Statistika STIS*
Jln. Otto Iskandardinnata No 64C, Bidara Cina, Jatinegara, Jakarta, 13330, Indonesia

[10] *System Dynamics Centre SDC*
Jln. Mercurius C4, Bogor, 16680, Indonesia

*Corresponding author's e-mail: ** ***bagussumargo@unj.ac.id***

ABSTRACT

Article History:

Received: 5th January 2025
Revised: 18th February 2025
Accepted: 3rd April 2025
Published: 1st July 2025

Keywords:

Clustering;
DBSCAN;
KNN-Distance graph;
Stunting prevalence;
Typology.

Stunting, a condition where children are malnourished for a long period, causes growth failure in children. West Java, Central Java, and East Java are the 3 provinces with the highest prevalence of stunting in 2021. This study aims to group districts/cities in these provinces based on factors that influence stunting using the DBSCAN method (there has been no previous research using this method for this case), so the typology of stunting prevalence is implied. The group results can be valuable input for policy priorities in overcoming stunting. The study used the DBSCAN (Density-Based Spatial Clustering of Application with Noise) method, which can also detect noises (outliers). The determination of eps and MinPts is based on the average value of the distance from each data to its closest neighbor. The distance obtained then was used in the KNN algorithm to determine eps and MinPts parameters. Clustering is done using standardized data and DBSCAN parameters obtained from the k-dist plot, eps is 1.92, and MinPts is 2. The validation test used is the silhouette coefficient to determine the goodness of the cluster results. The clustering results show that there are 2 clusters and 1 noise that have special characteristics related to factors that influence the prevalence of stunting. Cluster 1 consisted of 97 districts/cities and was characterized by a high percentage of infants under 6 months receiving exclusive breastfeeding and the lowest average per capita household expenditure. Cluster 2 (Bekasi City and Depok City) was characterized by the lowest percentage of households with proper health facilities and infants aged 0-59 months receiving complete immunization. The noise (high stunting prevalence) in Bandung City is characterized by the lowest percentage of households having proper sanitation.

How to cite this article:

B. Sumargo, Kadir, D. Safariza, M. Asikin, D. Siregar, N. N. Sari, D. Umbara, R. Hilmianto, R. Kurniawan and I. Firmansyah., "IMPLEMENTATION OF THE DBSCAN ALGORITHM FOR CLUSTERING STUNTING PREVALENCE TYPOLOGY IN WEST JAVA, CENTRAL JAVA, AND EAST JAVA REGIONS," *BAREKENG: J. Math. & App.*, vol. 19, no. 3, pp. 1779-1790, September, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Stunting is a chronic nutrition problem caused by medical conditions [1]. Stunting is when a child is malnourished for a long period, causing a growth failure [2], [3]. It is caused by the lack of diet intake that meets the nutritional needs of the fetus and the early growth period of a child. Stunting typically affects young children (below five years old) [4]. That malnourishment causes the child not to grow optimally like other normal children at the same age. Children with stunting as toddlers may face inhibited growth and will not achieve full height. Stunting causes physical harm to health and impacts children's intelligence. Their brains do not reach their maximum cognitive capabilities, resulting in permanent and severe cognitive damage [5]. In addition, stunting also causes reduced productivity, less income per person, higher chances of being in poverty, more potential for pregnancy complications and breathing difficulties during labor, and an elevated risk of chronic conditions such as obesity, diabetes, heart disease, stroke, high blood pressure, and cancer [6].

The main contributors to stunting, among others, are the lack of breast milk intake exclusively during 6 months since birth, family economic status, mother's body height, education level of the parent, and poor access to decent sanitation and clean water [7]. Other contributing factors are a history of infectious disease, immunization history, and protein intake [8]. Stunting is indirectly influenced by external factors, such as family socio-economic conditions, the mother's education level, and the family's income level [9]. Stunting can be prevented during the first 1.000 days after birth by providing extra food and fortifying food with iron. Additional ways to avoid stunting include ensuring adequate nutrition during pregnancy, practicing exclusive breastfeeding for newborns up to 2 years old, monitoring child growth and development, managing breast milk, introducing complementary foods, and keeping the environment clean [4].

Based on the World Health Organization (WHO), a country is said to have a stunting problem if the number of cases reaches above 20%. Meanwhile, in Indonesia, the prevalence rate of stunting reaches above normal by 20% [1]. In 2020, Indonesia became a country with the second highest stunting prevalence rate in Southeast Asia [10], [11]. In 2021, the prevalence of stunting in Indonesia was 24.4 %, our equivalent of 1.341.678 cases. The figure is still higher than WHO's threshold. In 2021, the Indonesian Nutritional Status Survey revealed that the prevalence rate of stunting in West Java, Central Java, and East Java, respectively, was 24.5%, 20.9%, and 23.5% [1]. The three provinces are among the five provinces with the highest stunting prevalence rates [12]. In addressing stunting, knowledge. and understanding its characteristics are very important. One way to overcome stunting is using the cluster method. Clustering is the process of grouping a large amount of data into several groups according to their respective characteristics [13].

Generally, clustering can assist in recognizing stunting characteristics in the three provinces. According to [14], the clustering methods mainly used for stunting are the non-hierarchy method (partition), hierarchy, density-based method, and grid-based method. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a well-suited clustering technique for spatial datasets [15]. Among the various types of clustering algorithms, the DBSCAN method is significantly more effective in finding clusters with changing shapes. It can also find clusters with uncertain shapes and does not require a predetermined number of clusters [16], [17]. Also, the DBSCAN algorithm can detect noise data that does not belong to any cluster [18]. DBSCAN achieves increased homogeneity and diversity by conducting personalized clustering on datasets with varying densities [17]. Previous research conducted in 2024 compares DBSCAN with several other clustering methods, such as K-Means and X-Mean, based on cluster validity using the Davies-Bouldin Index (DBI) method. The result of this study shows that DBSCAN produces the best validity clusters compared to the other two methods [19].

In the clustering algorithm, a distance function is needed to measure whether or not two particular objects are similar. Some distance functions usually used are Euclidian distance and Manhattan distance. Based on the [20] experiment, Euclidian distance performed better than other methods, such as Manhattan and Minkowski distance. After clustering, the validation is conducted to evaluate the clustering results. It is performed to assess the quality and ensure whether the results have represented the general population or not, as well as to avoid the difference in the clustering results. The silhouette coefficient is one method used to test the validity of clustering results. It has metric similarity between objects in the data set as its advantage. The value of the silhouette coefficient is between -1 and 1.

However, no one has used the DBSCAN method to cluster potentially stunted regions. DBSCAN is one of the best clustering algorithms for datasets with unknown structures, varying densities, and noise. Its ability to find arbitrarily shaped clusters without requiring k makes it more flexible than k-means and hierarchical clustering. However, choosing the right parameters (ε and *minPts*) is crucial for optimal performance. As a result of DBSCAN clustering, there may be districts or cities that are known for having the fewest factors that cause stunting within each cluster, along with the noise. The goal of this study is to sort areas of West Java, Central Java, and East Java Provinces into groups based on how common stunting is. This will help researchers come up with policy priorities that the governments of these provinces should focus on.

2. RESEARCH METHODS

2.1 Data

Data used in this study is secondary data published by BPS-Statistics Indonesia and the extension offices of the Ministry of Agriculture in West Java, Central Java, and East Java Provinces. The data analyzed consists of 27 cities/municipalities in West Java, 35 cities/municipalities in Central Java, and 38 cities/municipalities in East Java in 2021. Therefore, 100 cities/municipalities were analyzed in the study.

2.2 Research Variables

Variables used in the study are factors considered to influence the prevalence of stunting in the three provinces in 2021, as follows:

X_1 : The percentage of infants under six months who received breast milk exclusively

X_2 : The average household per capita expenditure

X_3 : The percentage of households with decent sanitation

X_4 : Percentage of households with access to decent health facilities

X_5 : Percentage of children aged 0-59 months received complete immunization

2.3 Data Standardization

Standardization is crucial for DBSCAN because it ensures fair distance calculations, unbiased clustering, and optimal parameter selection. Without it, DBSCAN might fail to detect meaningful clusters or misclassify points due to scale imbalances. Data standardization is a transformation process to equalize the scale of features with different units or ranges so that each feature contributes proportionally to the analysis without the dominance of certain features [21]. Data standardization can be done using a Z-score, which is formulated as follows:

$$Z_i = \frac{x_i - \bar{x}}{s} \tag{1}$$

where:

Z_i : Standardized value (z-score) of the i-th data

x_i : The value of the i-th observation data

$\bar{x}$: Mean value of observation data

s : Standard deviation of observation data

2.4 Cluster Analysis

Cluster analysis is a technique used to identify similar objects or individuals based on several criteria. The main goal of cluster analysis is to group related individuals or objects into smaller, mutually exclusive sets. A cluster must exhibit high internal homogeneity (within the cluster) and high external heterogeneity (between clusters) [22].

2.5 Euclidean Distance

One of the distance measurement formulas for data with numerical attributes is the Euclidean distance. Euclidean distance is the distance between two data points (i, j) with n numerical attributes, represented as $\boldsymbol{i} = x_{i1}, x_{i2}, \dots, x_{in}$ and $\boldsymbol{j} = x_{j1}, x_{j2}, \dots, x_{jn}$ in an n-dimensional space (R^n) [23]:

$$d(i,j) = \sqrt{\left(x_{i1} - x_{j1}\right)^2 + \left(x_{i2} - x_{j2}\right)^2 + \cdots + \left(x_{in} - x_{jn}\right)^2} \tag{2}$$

2.6 Density-Based Spatial Clustering Algorithm with Noise (DBSCAN)

DBSCAN is a clustering algorithm based on the density of data [23]. Unlike partition-based methods (K-Means, K-Medoids) requiring predetermined cluster numbers, this technique naturally handles noise points while effectively capturing clusters with complex, non-spherical configurations [24]. The concept of density in DBSCAN refers to the number of data points within the radius defined by MinPts (the minimum number of data points within the eps radius). Data points that fall within this radius are considered to have the desired density, including the data point itself. This density concept results in three types of statuses for each data point: core, border, and noise. A data point is classified as a core if the number of neighboring points within the eps radius, including the point itself, is greater than or equal to MinPts. For data points with fewer than MinPts neighbors within the eps radius but where the neighboring points are cores due to their presence, the data point is classified as a border. If a data point has fewer than MinPts neighbors and none of its neighbors are cores, it is classified as noise [25]. A cluster is defined as a set of connected points with maximum density. DBSCAN determines the number of clusters to be formed on its own but requires two additional inputs: (1) MinPts: the minimum number of items in a cluster, and (2) eps: the distance value between items that forms the basis for the neighborhood of a data point. To determine the optimal values for eps and MinPts, one can construct a k-distance graph by observing the shift in the value of eps from varying k values. The point at which a sharp change occurs in the k-distance graph will be used as eps, and the value of k will be used as MinPts [26].

2.7 Silhoutte Coefficient

The quality of a clustering result can be evaluated using the silhouette coefficient, which measures how well data points are grouped within their assigned clusters (intra-cluster) and how distinctly they are separated from other clusters (inter-cluster) [27]. The intra-cluster distance is the average distance between a data point and other points within the same cluster. In contrast, the inter-cluster distance is the average distance between a data point and points in a different cluster [28]. A higher silhouette coefficient indicates better clustering quality, reflecting a greater difference between intra-cluster compactness and inter-cluster separation. The steps in calculating the silhouette coefficient are as follows [29]:

a. Determine the average distance of an object, such as the i-th object, to all other objects within the same cluster with the Equation (3):

$$a(i) = \frac{1}{n_A - 1} \sum_{j \in A, j \neq i}^{n_A} d(i,j) \tag{3}$$

where:

$a(i)$: Intra-cluster distance.

n_A : The total number of points in cluster $\boldsymbol{A}$.

$d(i,j)$: The distance between point i and point j.

b. Determine the average distance from the i-th object to all objects in other clusters. For instance, let C represent all clusters except A. The average distance can then be calculated using **Equation (4)**:

$$b(i) = \min_{C \neq A} \left(\frac{1}{n_C} \sum_{j \in C}^{n_C} d(i,j) \right) \tag{4}$$

where:

$b(i)$: Inter-cluster distance.

n_C : The total number of points in cluster C.

$d(i,j)$: The distance between point i and point j.

c. After calculating the intra-cluster distance and inter-cluster distance, the silhouette value can be calculated with the **Equation (5)**:

$$s(i) = \frac{b(i) - a(i)}{\max\left(a(i); b(i)\right)} \tag{5}$$

d. After that, calculate Silhoutte Coefficient (SC) with the **Equation (6)**:

$$SC = \frac{1}{n} \sum_{i=1}^{n} s(i) \tag{6}$$

where:

SC : Silhoutte Coefficient with a value range between $-1 < SC < 1$.

The SC value indicates the degree of similarity between objects within a cluster. When the SC value is close to 1, it suggests that the clustering results are of higher quality. The measurement criteria of the Silhoutte Coefficient are given in **Table 1**.

Table 1. Silhouette Coefficient Measurement Criteria

SC Value	Criterion
0.71 – 1.00	Strong structure
0.51 – 0.70	Good structure
0.26 – 0.50	Weak structure
≤0.25	Poor structure

2.8 Stage of Analysis

a. Completing descriptive statistics.

b. Conducting standardization factor data influencing stunting in the three provinces using z-score.

c. Determining the initial core dot randomly.

d. Determining Euclidian distance by computing the nearest distance from the point of the data to a centroid.

e. Determine eps and MinPts by looking at the plot of the KNN-distance graph.

f. Inspecting the number of dots in a radius of $eps \geq MinPts$

g. Ensure that all dots have been processed. In case any dots have not been processed, the fourth stage is repeated. If all dots have been processed, then continue to the next stage.

h. Conducting cluster validation based on the results of clustering using the *Silhouette Coefficient* to determine a good cluster.

i. Analyze clustering results to know which cities/municipalities form each cluster and have contributed to the stunting prevalence.

3. RESULTS AND DISCUSSION

3.1 Results

All the equation numbers must be aligned. Data from 100 cities/municipalities in the three provinces are used as the object of the study of the stunting prevalence. We give a summary of the variable that we used **Table 2**. All variables are standardized into z-scores to reduce the variation and eliminate the effect of the measurement scale differences between variables. After standardization, a dot, Bogor Regency, was selected randomly, and then, from the initial dot, we calculated the Euclidian distance between Bogor Regency and Sukabumi.

The distance obtained then was used in the KNN algorithm to determine eps and MinPts parameters. The determination of eps and MinPts is based on the average value of the distance from each data to its closest neighbor. The plot of the known-distance graph is given in **Figure 1**, where the plot is used to determine the optimal value of eps using a MinPts value of 2.

As given in **Figure 1**, using $k = 2$, the optimum value of eps obtained is between 1.55 and 1.92 because sharp change occurs along the k-distance curve. Therefore, this interval was selected. With the optimum value for eps is 1.92 and the value of MinPts is 2, we obtained the silhouette coefficient value of 0.28.

Table 2. Variable Description Used in Analysis

Variable	Minimum	Maximum	Mean
Percentage of infants under six months receiving breast milk exclusively (X_1)	19.24	100	40.6515
Average household per capita expenditure (X_2)	663,069	2,488,463	1,167,226
Percentage of households with decent sanitation (X_3)	39.44	98.07	80.00
Percentage of households with access to decent health facilities (X_4)	43.14	100	14.77
Percentage of children aged 0-59 months received complete immunization (X_5)	0	100	88.94

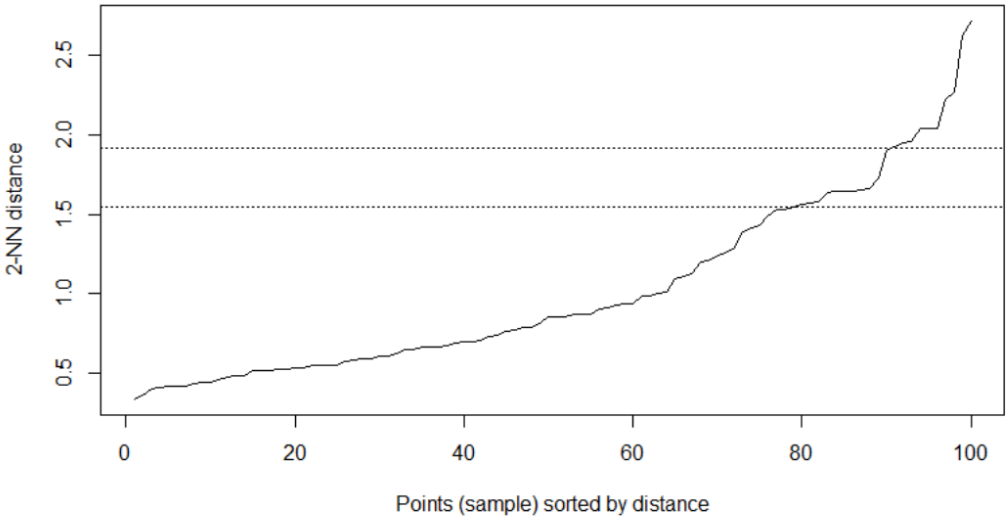

Figure 1. K-Dist KNN Plot, $k = 2$

Clustering is done using standardized data and DBSCAN parameters obtained from the k-dist plot, eps is 1.92, and MinPts is 2. The clustering results are obtained by applying the KNN algorithm using the Euclidean distance, as shown in **Figure 2**.

The cluster results in **Figure 2** show that there are two clusters. The first cluster consists of 97 cities/municipalities, and the second consists of 2 cities/municipalities. Apart from that, there is also noise, a city/municipality that does not belong to any cluster formed. The cluster results in this study show different characteristics between clusters.

Figure 2. Clustering Results

After the cluster is formed, the next step is to validate the cluster results using the silhouette coefficient. Here is the calculation of the silhouette coefficient in Bogor Regency, $S(i)$, is as follows:

$$S(i) = \frac{b(i)-a(i)}{\mathrm{Max}(a(i),b(i))}$$

$$= \frac{4{,}318-4{,}324}{Max[4{,}324\ ,4{,}318]}$$

$$= \frac{4{,}318-4{,}324}{4{,}324} = -0.001$$

Once the coefficient of Bogor Regency is calculated, the next step is to calculate the silhouette coefficient of all other data in each cluster and noise. The silhouette coefficient for all data in each cluster and noise is presented in **Figure 3**.

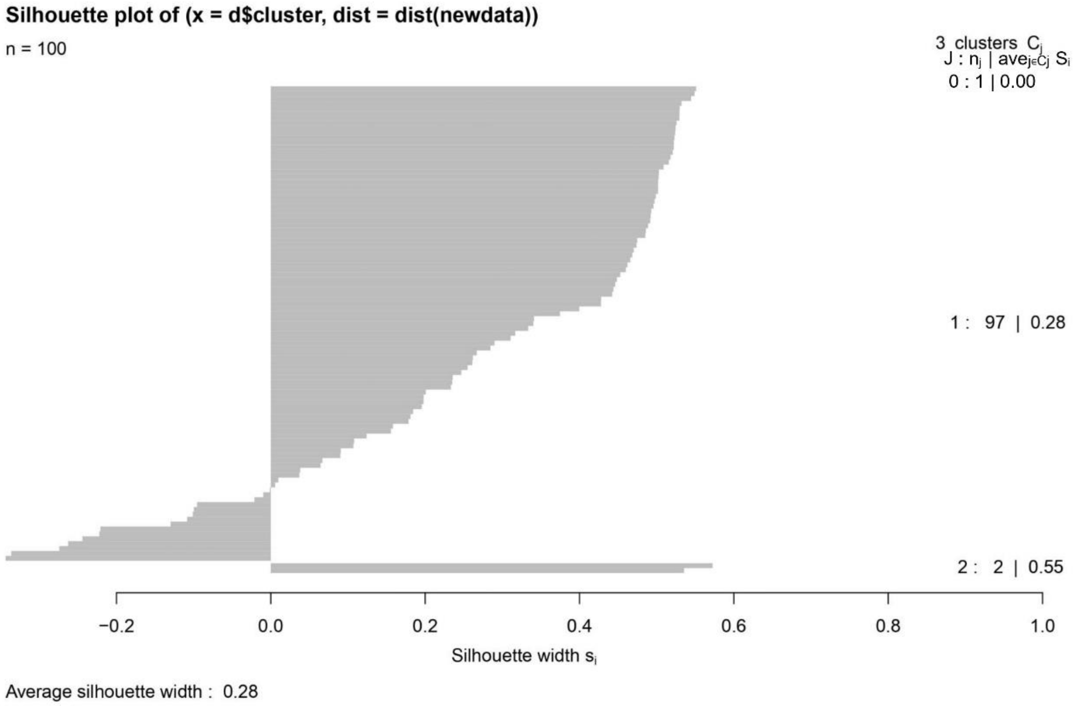

Figure 3. Silhouette Coefficient Plot

Based on the Silhouette Coefficient Plot, it was obtained that the average silhouette coefficient value is 0,28 meaning that the cluster has a weak structure.

Each formed cluster has different characteristics that can be seen by averaging the values of each variable in the cluster. The aim is to determine the uniqueness and differences between each formed cluster. The results of the calculation of the average for each variable are shown in **Table 3**. The characteristics of each cluster (contributing factors of stunting) based on the average of all variables in each cluster and noise are presented in **Table 3**.

In cluster 1, the percentage of infants under six months who received breast milk exclusively is 39.883%. That value is still under the Ministry of Health's Health's target of 45%. The average household per capita expenditure is as much as Rp1,132,441 monthly or belongs to the expenditure group below Rp 1,500,000 on average monthly or low category [30]. The percentage of households with decent sanitation is around 79.965%. The figure is still below the SDG target of 100 % of total households. The rate of households with access to decent health facilities is 14.74%, which is also still below the target of 51%. Moreover, the percentage of children aged 0-59 months who received complete immunization is 63.73%, which is also far below the Ministry of Health's target of 93% in 2021.

In cluster 2, the percentage of infants under six months who received breast milk exclusively is 61.95%. That value is already above the Ministry of Health’s target of 45%. The average household per capita expenditure is as much as Rp 2,396,702 monthly or belongs to the expenditure group between Rp 1,500,000 and Rp 2,500,000 on average monthly or medium category [27]. he percentage of households with decent sanitation is around 97.3%, close to the SDG target of 100 % of total households. The percentage of households with access to decent health facilities is 13.26%, which is still far below the target of 51%. Moreover, the percentage of children aged 0-59 months who received complete immunization is 58.62%, which is also far below the Ministry of Health’s target of 93% in 2021.

In noise, the percentage of infants under six months who received breast milk exclusively is 72.64%. That value is already above the Ministry of Health’s target (2021b). The average household per capita expenditure is as much as Rp 2,082,374 monthly or belongs to the expenditure group between Rp 1,500,000 and Rp 2,500,000 on average monthly or medium category [30]. The percentage of households with decent sanitation is around 48.9%, close to the SDG target. The percentage of households with access to decent health facilities is 19.95%, which is still far below the target. Moreover, the percentage of children aged 0-59 months who received complete immunization is 66.31%, which is also far below the Ministry of Health's target in 2021.

The highest prevalence of stunting is in noise, which is 23%, followed by Cluster 1, with an average prevalence of 21%, and Cluster 2, with the lowest average cluster prevalence of 19%. So, the priority is the noise and cluster 1 municipalities because it has a high stunting prevalence.

Tabel 2. The Average of Stunting Contributing

Variable	Cluster 1	Cluster 2	Noise
Percentage of infants under six months receiving breast milk exclusively (X_1)	39.883	61.95	72.64
Average household per capita expenditure (X_2)	1,132,441	2,396,702	2,082,374
Percentage of households with decent sanitation (X_3)	79.965	97.3	48.9
Percentage of households with access to decent health facilities (X_4)	14.743	13.26	19.95
Percentage of children aged 0-59 months received complete immunization (X_5)	63.726	58.615	66.31
Stunting prevalence	**21.488**	**13.05**	**26.4**

Figure 4 shows that cluster 1, which is green in color, consists of 97 municipalities, of which there are 24 municipalities in West Java, 35 municipalities in Central Java, and municipalities in East Java. Cluster 2, which is in blue, only consists of two municipalities, while the noise in red only consists of one municipality. **Figure 4** shows the clustering results of municipalities by province.

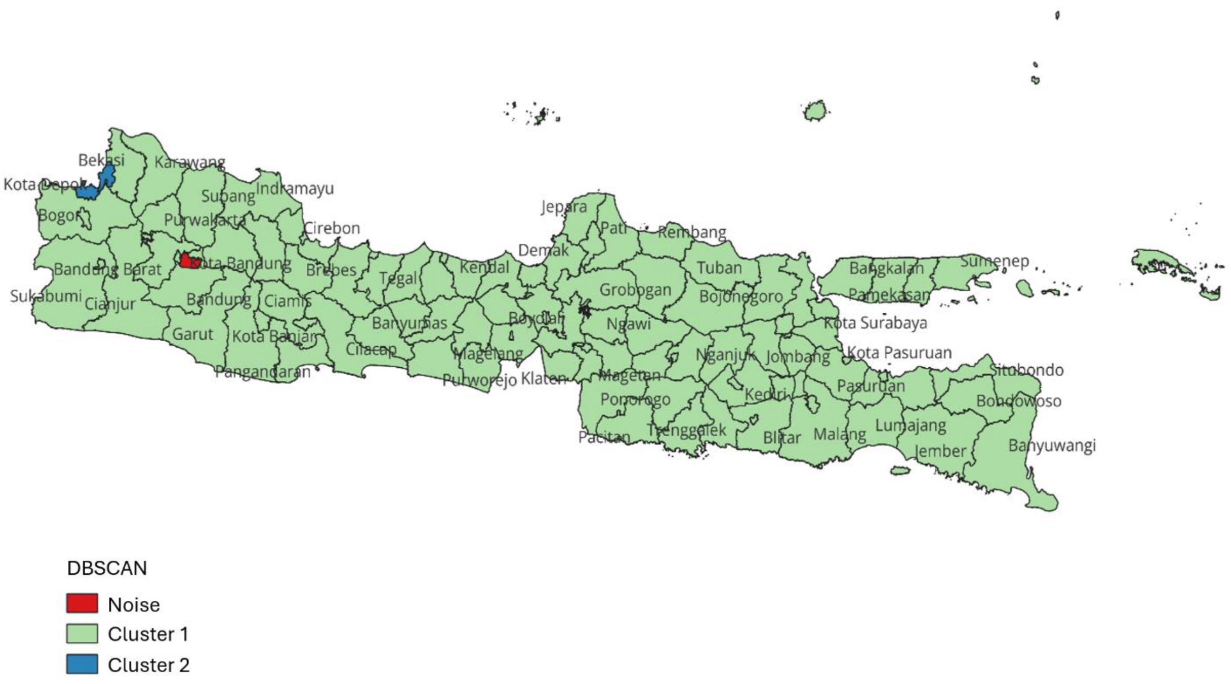

Figure 4. Clustering Results of Municipalities

3.2 Discussion

Clustering results give information on the typology of stunting prevalence, which is the dominant or stunting factor high in a region. Cluster 1 consists of 97 municipalities with the lowest percentage of infants under six months who received breast milk exclusively and household expenditure per capita. Cluster 2 consists of Depok and Bekasi, which are characterized by the lowest households with access to decent health facilities and the percentage of children aged 0-59 months who received complete immunization. The lowest percentage of households with decent sanitation characterizes the noise, which has a high prevalence of stunting in Bandung City. Our findings show that all factors took place in all municipalities in West Java. At the same time, factor influencers of stunting in East Java and Central Java are characterized by the lowest percentage of infants under six months receiving breast milk exclusively and the average household per capita expenditure.

4. CONCLUSIONS

Stunting is a complex public health issue influenced by multiple factors, including socioeconomic conditions, maternal health, nutrition, sanitation, and healthcare access. We can use DBSCAN clustering to group regions or populations with similar stunting patterns, helping policymakers effectively target interventions. DBSCAN is effective in the clustering areas based on stunting prevalence. Its limitations can impact its performance, such as the fact that DBSCAN assumes clusters have similar densities. However, stunting prevalence may vary widely between urban and rural areas; DBSCAN assumes clusters have identical densities. However, stunting prevalence may vary widely between urban and rural areas; DBSCAN performs best in low-dimensional spaces. When using multiple factors (e.g., nutrition, healthcare, income, sanitation), distance calculations may become unreliable due to the curse of dimensionality. DBSCAN is a static clustering method that does not naturally handle time-series data. Despite these limitations, DBSCAN remains a powerful tool for clustering stunting prevalence data. Future research should focus on Hybrid approaches, combining DBSCAN with HDBSCAN or machine learning models for better accuracy; Integrating geospatial data, using GIS-based clustering to identify high-risk regions; and developing dynamic models, incorporating time-series analysis to track stunting trends. The number of clusters formed is 2 and 1 noise with a silhouette coefficient of 0.28. The typology of stunting factors in all municipalities in West Java, Central Java, and East Java is characterized by the lowest percentage of infants under six months receiving breast milk exclusively and the average household per capita expenditure. DBSCAN heavily depends on the correct choice of ε (epsilon) and minPts. By addressing these challenges, DBSCAN can become an even more effective tool for policymakers in identifying and mitigating stunting prevalence in different regions, particularly those related to risk factors, the lowest percentage of infants under six months receiving breast milk exclusively, and the average household per capita expenditure. This research can be improved by adding more factors influencing the stunting prevalence and using different methodological approaches in identifying the typology.

ACKNOWLEDGMENT

Our gratitude goes to the authors for their extraordinary cooperation so that this article can be completed and hopefully be helpful. Our gratitude goes to the leadership of the Universitas Negeri Jakarta, which has funded this research by the Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta, Number: 82/SPK PENELITIAN/FMIPA/2025, March 19, 2025.

REFERENCES

[1] Kemenkes RI, *BUKU SAKU HASIL STUDI STATUS GIZI INDONESIA (SSGI) TINGKAT NASIONAL, PROVINSI, DAN KABUPATEN/KOTA TAHUN 2021*. Jakarta, 2021. doi: 10.36805/bi.v2i1.301.

[2] R. R. Akbar, W. Kartika, and M. Khairunnisa, "THE EFFECT OF STUNTING ON CHILD GROWTH AND DEVELOPMENt," *Sci. J.*, vol. 2, no. 4, pp. 153–160, 2023, doi: https://doi.org/10.56260/sciena.v2i4.118.

[3] C. R. Montenegro *et al.*, "THE PEDIATRIC GLOBAL BURDEN OF STUNTING: FOCUS ON LATIN AMERICA," *Lifestyle Med.*, vol. 3, no. 3, pp. 1–11, 2022, doi: https://doi.org/10.1002/lim2.67.

[4] S. Wulandari, "CLUSTERING INDONESIAN PROVINCES ON PREVALENCE OF STUNTING TODDLERS USING AGGLOMERATIVE HIERARCHICAL CLUSTERING," *Fakt. Exacta*, vol. 16, no. 2, 2023, doi: https://doi.org/10.30998/faktorexacta.v16i2.17186.

[5] A. J. Pitoyo, A. Saputri, R. E. Agustina, and T. Handayani, "ANALYSIS OF DETERMINAN OF STUNTING PREVALENCE AMONG STUNTED TODDLERS IN INDONESIA," *Populasi*, vol. 30, no. 1, p. 36, 2022, doi: https://doi.org/10.22146/jp.75796.

[6] Azizah Andzar Ridwanah, Hario Megatsari, Agung Dwi Laksono, and Mursyidul Ibad, "FACTORS RELATED TO STUNTED IN EAST JAVA PROVINCE IN 2019: AN ECOLOGICAL ANALYSIS," *Med. Leg. Updat.*, vol. 21, no. 2, pp. 230–235, 2021, doi: https://doi.org/10.37506/mlu.v21i2.2678.

[7] T. Beal, A. Tumilowicz, A. Sutrisna, D. Izwardy, and L. M. Neufeld, "A REVIEW OF CHILD STUNTING DETERMINANTS IN INDONESIA," *Matern. Child Nutr.*, vol. 14, no. 4, pp. 1–10, 2018, doi: https://doi.org/10.1111/mcn.12617.

[8] L. I. P. Ariati, "FAKTOR-FAKTOR RESIKO PENYEBAB TERJADINYA STUNTING PADA BALITA USIA 23-59

BULAN," *OKSITOSIN J. Ilm. Kebidanan*, vol. 6, no. 1, pp. 28–37, 2019, doi: https://doi.org/10.35316/oksitosin.v6i1.341.

[9] F. O. Aridiyah, N. Rohmawati, and M. Ririanty, "FAKTOR-FAKTOR YANG MEMPENGARUHI KEJADIAN STUNTING PADA ANAK BALITA DI WILAYAH PEDESAAN DAN PERKOTAAN (THE FACTORS AFFECTING STUNTING ON TODDLERS IN RURAL AND URBAN AREAS)," *e-Jurnal Pustaka Kesehat.*, vol. 3, no. 1, 2015, doi: 10.1007/s11746-013-2339-4.

[10] J. Jumhur, "THE EFFECT OF ECONOMIC GROWTH AND POVERTY ON STUNTING IN INDONESIA," *J. Perspekt. Pembiayaan dan Pembang. Drh.*, vol. 11, no. 6, pp. 433–448, 2024, doi: https://doi.org/10.22437/ppd.v11i6.26871.

[11] H. Hatijar, "The Incidence of Stunting in Infants and Toddlers," *J. Ilm. Kesehat. Sandi Husada*, vol. 12, no. 1, pp. 224–229, 2023, doi: 10.35816/jiskh.v12i1.1019.

[12] Taufik Hidayat, Mohamad Jajuli, and Susilawati, "CLUSTERING DAERAH RAWAN STUNTING DI JAWA BARAT MENGGUNAKAN ALGORITMA K-MEANS," *INFOTECH J. Inform. Teknol.*, vol. 4, no. 2, pp. 137–146, 2023, doi: https://doi.org/10.37373/infotech.v4i2.642.

[13] S. S. Nagari and L. Inayati, "IMPLEMENTATION OF CLUSTERING USING K-MEANS METHOD TO DETERMINE NUTRITIONAL STATUS," *J. Biometrika dan Kependud.*, vol. 9, no. 1, pp. 62–68, 2020, doi: https://doi.org/10.20473/jbk.v9i1.2020.62-68.

[14] M. A. Robbani, G. Firmansyah, and A. M. Widodo, "CLUSTERING OF CHILD STUNTING DATA IN TANGERANG REGENCY USING COMPARISON OF K-MEANS , HIERARCHICAL CLUSTERING AND DBSCAN METHODS," vol. 2, no. 2015, pp. 3105–3112, 2024.

[15] M. Bariklana and A. Fauzan, "IMPLEMENTATION OF THE DBSCAN METHOD FOR CLUSTER MAPPING OF EARTHQUAKE SPREAD LOCATION," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 2, pp. 0867–0878, 2023, doi: https://doi.org/10.30598/barekengvol17iss2pp0867-0878.

[16] T. D. Harjanto, A. Vatresia, and R. Faurina, "ANALISIS PENETAPAN SKALA PRIORITAS PENANGANAN BALITA STUNTING MENGGUNAKAN METODE DBSCAN CLUSTERING," *J. Rekursif*, vol. 9, no. 1, pp. 30–42, 2021.doi: https://doi.org/10.33369/rekursif.v9i1.14982

[17] N. P. Sutramiani, I. M. T. Arthana, P. F. Lampung, S. Aurelia, M. Fauzi, and I. W. A. S. Darma, "THE PERFORMANCE COMPARISON OF DBSCAN AND K-MEANS CLUSTERING FOR MSMES GROUPING BASED ON ASSET VALUE AND TURNOVER," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 10, no. 1, pp. 13–24, 2024, doi: https://doi.org/10.20473/jisebi.10.1.13-24.

[18] B. Tan, M. Steinbach, A. Karpatne, and V. Kumar, *CLUSTER ANALYSIS: BASIC CONCEPT AND ALGORITHMS*. New York: Pearson Education, 2019.

[19] V. Wulandari, Y. Syarif, Z. Alfian, M. A. Althof, and M. Mufidah, "COMPARISON OF DENSITY-BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE (DBSCAN), K-MEANS AND X-MEANS ALGORITHMS ON SHOPPING TRENDS DATA," *IJATIS Indones. J. Appl. Technol. Innov. Sci.*, vol. 1, no. 1, pp. 1–8, 2024, doi: https://doi.org/10.57152/ijatis.v1i1.1135.

[20] M. Nishom, "PERBANDINGAN AKURASI EUCLIDEAN DISTANCE, MINKOWSKI DISTANCE, DAN MANHATTAN DISTANCE PADA ALGORITMA K-MEANS CLUSTERING BERBASIS CHI-SQUARE," *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 20–24, 2019, doi: https://doi.org/10.30591/jpit.v4i1.1253.

[21] C. Andrade, "Z SCORES, STANDARD SCORES, AND COMPOSITE TEST SCORES EXPLAINED," *Indian J. Psychol. Med.*, vol. 43, no. 6, pp. 555–557, 2021, doi: https://doi.org/10.1177/02537176211046525.

[22] A. Migliore and C. Rossi-Lamastra, "CLUSTER ANALYSIS: GROUPING WORKERS BY WORK LOCATION CHOICE," *Methodol. Approaches Work. Res. Manag.*, pp. 95–107, 2023, doi: https://doi.org/10.1201/9781003289845-7.

[23] Suyanto, *DATA MINING UNTUK KLASIFIKASI DAN KLASTERISASI DATA*, Penerbit I. Bandung, 2019.

[24] I. Syahzaqi, M. Effendi, H. Rahmawati, H. Kuswanto, and S. Sediono, "GROUPING PROVINCES IN INDONESIA BASED ON THE NUMBER OF VILLAGES AFFECTED BY ENVIROMENTAL POLLUTION WITH K-MEDOIDS, FUZZY C-MEANS, AND DBSCAN," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 18, no. 2, pp. 0923–0936, 2024, doi: https://doi.org/10.30598/barekengvol18iss2pp0923-0936.

[25] P. Silitonga, "ANALISIS POLA PENYEBARAN PENYAKIT PASIEN PENGGUNA BADAN PENYELENGGARA JAMINAN SOSIAL (BPJS) KESEHATAN DENGAN MENGGUNAKAN METODE DBSCAN CLUSTERING," *TIMES*, vol. V, no. 1, pp. 36–39, 2016.doi: https://doi.org/10.51351/jtm.5.1.2016482

[26] M. S. Amir, "HOW TO DETERMINE EPSILON AND MINPTS PARAMETERS OF DBSCAN CLUSTERING," Sefidian Academy, [Online]. [Online]. Available: https://www.sefidian.com/2022/12/18/how-to-determine-epsilon-and-minpts-parameters-of-dbscan-clustering/

[27] A. Vysala and D. J. Gomes, "EVALUATING AND VALIDATING CLUSTER RESULTS," pp. 37–47, 2020, doi: https://doi.org/10.5121/csit.2020.100904.

[28] N. Kaoungku, K. Suksut, R. Chanklan, K. Kerdprasop, and N. Kerdprasop, "THE SILHOUETTE WIDTH CRITERION FOR CLUSTERING AND ASSOCIATION MINING TO SELECT IMAGE FEATURES," *Int. J. Mach. Learn. Comput.*, vol. 8, no. 1, pp. 69–73, 2018, doi: https://doi.org/10.18178/ijmlc.2018.8.1.665.

[29] H. Řezanková, "DIFFERENT APPROACHES TO THE SILHOUETTE COEFFICIENT CALCULATION IN CLUSTER EVALUATION," *21st Int. Sci. Conf. AMSE*, no. September, pp. 1–10, 2018.

[30] L. S. Rakasiwi, "PENGARUH FAKTOR DEMOGRAFI DAN SOSIAL EKONOMI TERHADAP STATUS KESEHATAN INDIVIDU DI INDONESIA," *Kaji. Ekon. dan Keuang.*, vol. 5, no. 2, pp. 146–157, 2021, doi: https://doi.org/10.31685/kek.v5i2.1008.