Available online at http://scik.org

Commun. Math. Biol. Neurosci. 2024, 2024:10

https://doi.org/10.28919/cmbn/8365

ISSN: 2052-2541

CLUSTERING OF CHILDHOOD DIARRHEA DISEASES USING GAUSSIAN MIXTURE MODEL

DEFI YUSTI FAIDAH*, ASHILLA MAULA HUDZAIFA, RESA SEPTIANI PONTOH

Department of Statistics, Universitas Padjadjaran, Bandung 45363, Indonesia

Abstract: Diarrhea remains one of the leading causes of childhood morbidity and mortality worldwide, especially in developing countries. Diarrhea is a digestive disorder characterized by the need to defecate more often than usual. Diarrhea remains a major problem of public health in Indonesia, with 40% of cases occurring in children under five years old. Bandung City is one of the largest cities in West Java Province with the sixth highest cases of diarrhea in West Java. The high number of diarrhea cases in Bandung City is a serious concern in public health efforts, especially among young children. The incidence of diarrhea in children under five in Bandung City in 2022 was recorded as high as 6376 cases. Therefore, this study aims to cluster sub-districts in Bandung City based on factors that influence diarrhea cases. Five factors are used, namely the prevalence of diarrhea in children under five years old, clean and healthy living behaviour, healthy latrine facilities, population density, and many baby less than 6 months old are exclusively breastfed . Based on the results of clustering with the Gaussian mixture model, 5 regional groups were obtained based on similarity of characteristics with five factors. Cluster 2 are areas characterized by a high percentage of households with healthy latrines and healthy living behaviour. Cluster 5 are areas characterized by a high prevalence of diarrhea in children under five. The results of this clustering indicate the importance of interventions and strategies

*Corresponding author

E-mail address: defi.yusti@unpad.ac.id

Received November 21, 2023

by the government of Bandung city to prevent the increasing number of diarrhea cases among children under five. Therefore, it can provide better insight into the distribution of diarrhea cases to achieve the Sustainable Development Goals (SDGs) and improve the health status of children under five in Indonesia, particularly in Bandung City.

Keywords: childhood; clustering; diarrhea; Gaussian mixture model; SDGs.

2020 AMS Subject Classification: 92C50.

1. INTRODUCTION

Diarrhea remains one of the leading causes of childhood morbidity and mortality worldwide, especially in developing countries. Diarrhea is a digestive disorder characterized by the sufferer passing stools more frequently than usual. Diarrhea is caused by a number of things, for example Escherichia coli bacteria, viruses such as Rotavirus, parasites, food allergies, Crohn's disease, side effects of certain drugs, and others. Diarrhea can be classified based on the frequency and duration of bowel movements and the characteristics of the faeces [1]. The classification of diarrhea includes acute diarrhea lasting 14 days, persistent diarrhea lasting > 14 days, and chronic diarrhea lasting > 30 days [2][18]. Serious diarrhea can lead to malnutrition and in the most severe cases, death due to lack of salt and water in the body [19][20]. Indonesia as the fourth most populous country in the world is place to around 22 million children under five. The Health Ministry of the Republic of Indonesia indicates diarrhea as the leading cause of death in children under 5 years of age with a mortality rate of 10.7% in 2019 [3]. In addition to causing death, prolonged diarrhea can lead to malnutrition and stunting in children [4].

Diarrhea remains a major problem of public health in Indonesia with 40% of cases occurring in children under five years old. Bandung is one of the major cities with the sixth highest diarrhea cases in West Java. The high incidence of diarrhea in Bandung City is a serious concern in efforts to address public health, especially children under five. In 2022, the incidence of diarrhea in children under five in Bandung City was recorded as many as 6376 cases served. Figure 1 shows that Babakan Ciparay is the sub-district with the most diarrhea cases in Bandung City.

There are various factors that cause diarrheal diseases, one of which is hygiene problems including improper sanitation facilities [17]. Germs that cause diarrhea are easily spread from one

person to another through contaminated water, food or objects [5]. Food hygiene is associated with the development of diarrhea and malnutrition in low socioeconomic children [6]. In addition, limited and inadequate sanitation facilities are likely to have a bad level of hygiene and may increase the risk of diarrhea.

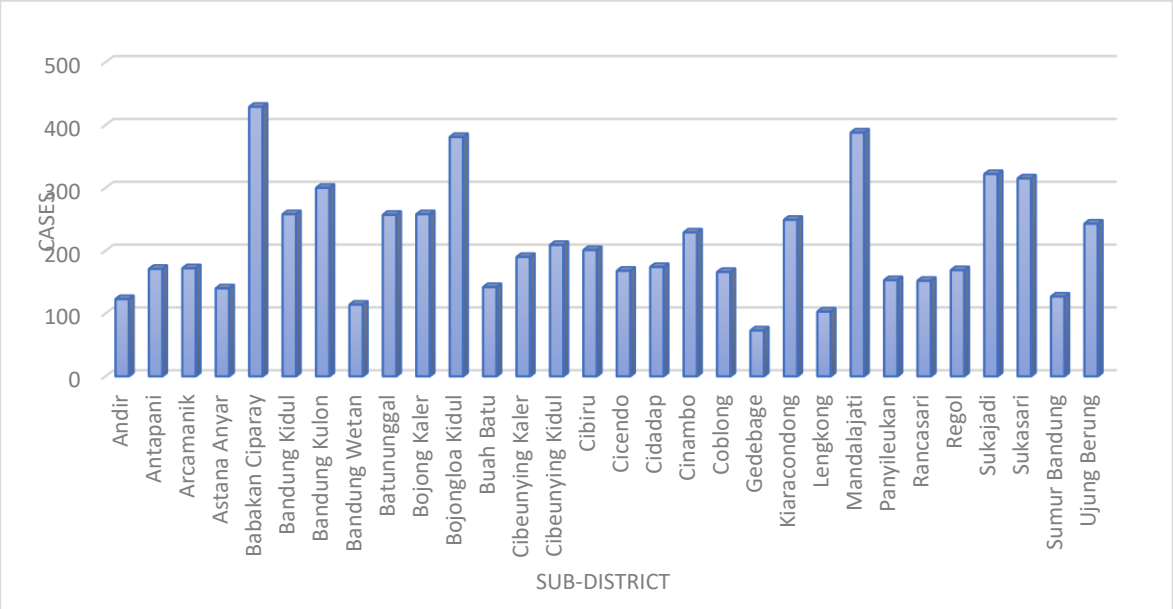

Figure 1. Many Cases of Diarrhea in Children Under Five in Kota Bandung.

Various programmes have been carried out to reduce the occurrence of diarrhea, one of which is the provision of clean water and sanitation in area that still difficult access to clean water. The provision of sanitation and clean water the goal 6 of SDGs [31]. However, morbidity and mortality rates from diarrhea remain high due to the high prevalence of contributing factors. Therefore, this study aims to cluster sub-districts in Bandung City based on the percentage of diarrhea prevalence, percentage of households with healthy latrine facilities, percentage of households with clean and healthy living behaviours, and population density per hectare as an effort to provide better insight into the distribution of diarrhea cases for the achievement of Sustainable Development Goals (SDGs).

The results of this study can show the importance interventions and efforts of the Bandung City government in preventing high cases of diarrhea and optimizing the improvement of health levels in children under five.

2. MATERIALS AND METHODS

2.1. Data

The data used for this study is secondary data obtained from the data portal and health profile of Bandung City includes several variables, namely the prevalence of diarrhae in children under five, the percentage of households with clean and healthy living behaviour, the percentage of households with healthy latrine facilities, the population density per-hectare, and the many baby less than 6 months old who are exclusively breastfed. This study used 30 observations, namely the number of sub-districts in Bandung City.

2.1.1. Prevalence of diarrhea in children under five

The prevalence of diarrhea in children under five years old is a record of the number of diarrhea cases identified and treated. The data used is the prevalence of diarrhea in children under five. In this study, the prevalence of diarrhea in children under five was measured in each sub-district in Bandung City.

$$Prevalence = \frac{Number\ of\ cases\ of\ diarrhoea\ in\ children\ under\ five}{Total\ cases\ in\ the\ region} \times 100\% \qquad (1)$$

2.1.2. Percentage of households with clean and healthy living behaviour

Households with Clean and Healthy Living Behaviour is all health behaviours which is done out of awareness and taking an active role in community health activities [7]. PHBS involves several elements, namely households, schools, workplaces, health facilities, and public places (Health Profile of Bandung City 2022). The data used in the study is the percentage of households identified as implementing clean and healthy living behaviours.

2.1.3. Percentage of households with healthy latrine facilities

Healthy latrines are proper sanitation facilities that are able to prevent themselves from various diseases. The criteria included in the operational definition for healthy latrines are a latrine building that is closed and has a non-slip floor, no odor and no visible dirt, septic tank distance ≥ 10 metres, available cleaning tools, and free from insects [8]. The data used in the study is the percentage of households that have healthy latrine facilities.

2.1.4. Population density of a sub-district per-hectare

Population density is the number of people per unit area (ha). Population density is an indicator to measure the population in an area [28]. This indicator is used to determine the population density per-hectare of an area.

2.1.5. Many babies less than 6 months who are exclusively breastfed.

Breast milk is the ideal food for baby up to 6 months old in terms of physical and psychological health [9]. Exclusive breastfeeding until the baby is 6 months old will influence the optimal development of the child's intelligence potential [10]. Exclusive breastfeeding in children before 6 months of age is very important to reduce the risk of developing various diseases and breast milk can accelerate recovery in sick children [29][30].

2.2. Data Standarization

Standardization is a technique used to transform data so that it has a mean equal to 0 and a standard deviation equal to 1. Standardization is used in data analysis when the observed variables have different scales or distributions [27]. Data values that have been standardized are notated as z, x is the actual data value, μ is the mean of the data, and σ is the standard deviation of the data.

$$z = \frac{x - \mu}{\sigma} \tag{2}$$

2.3. Variance Inflation Factor

Variance Inflation Factor (VIF) is a measure of the severity level of multicollinearity in multiple linear regression models involving more than one variable. Multicollinearity is a measure that refers to the comparison of variance when there is multicollinearity between predictor variables and variance when there is no multicollinearity. The formula for calculating VIF is as follows [11].

$$VIF_i = \frac{1}{1 - R_i^2} \tag{3}$$

R_i in the equation states the determination coefficient of the ith variable. The occurrence of multicollinearity in data based on the VIF value > 10 which indicates that the greater the VIF, the more serious the multicollinearity [12].

2.4. Gaussian Mixture Model

Gaussian Mixture Model (GMM) is a statistical model used to represent complex data distributions

by combining multiple Gaussian distributions [21][22]. In this model, data is considered to come from several different components of the Gaussian distribution. GMM is used in various fields including clustering [26], dimensionality reduction, data distribution modelling, image restoration, and others. In this research, the use of GMM aims to perform clustering.

GMM assumes that the resulting component of the Gaussian distribution is the number of clusters formed. The combination obtained from the mean and variance will represent each Gaussian. The purpose of clustering using GMM is to determine the model parameters (mean and matrix) that best fit the data [13]. The model used to perform clustering with respect to the geometry formed from components of Gaussian with different parameters [14] shown in table 1.

Table 1. Covariance matrix and geometric formed of Mclust in the Gaussian Mixture Model.

Symbol	Model	Volume	Geometry Shape	Orientation	Shape
EII	λI	Same	Same	-	Spherical
VII	$\lambda_k I$	Different	Same	-	Spherical
EEI	λA	Same	Same	Coordinate axes	Diagonal
VEI	$\lambda_k A$	Differrent	Same	Coordinate axes	Diagonal
EVI	λA_k	Same	Different	Coordinate axes	Diagonal
VVI	$\lambda_k A_k$	Different	Different	Coordinate axes	Diagonal
EEE	$\lambda D A D^T$	Same	Same	Identity	Ellipsoidal
EEV	$\lambda D_k A D_k^T$	Same	Same	Different	Ellipsoidal
VEV	$\lambda_k D_k A D_k^T$	Different	Same	Different	Ellipsoidal
VVV	$\lambda_k D_k A_k D_k^T$	Different	Different	Different	Ellipsoidal

The probability density function for a one-dimensional Gaussian distribution is:

$$f(X|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(X-\mu)^2}{2\sigma^2}} \quad (4)$$

μ and σ represent the mean and the standard deviation of the distribution. The probability density function for d-dimensional multivariate data is expressed as follows:

$$f(X|\mu,\sum) = \frac{1}{\sqrt{2\pi}|\sum|} e^{\left(-\frac{1}{2}(X-\mu)^T\sum^{-1}(X-\mu)\right)} \quad (5)$$

Where μ denotes the mean of the represented of distribution as a d-dimensional array, ∑

is the covariance matrix of X, T denotes the transpose vector, and -1 denotes the invers of the matrix [15]. To maximize the likelihood of data from GMM, the Expectation-Maximization (EM) algorithm can be used. The steps are as follows [13][23]:

1) Initialize the value of $\mu_k, \sigma_k,$ and π_k randomly for all clusters, where π is the mixture coefficient and k value is a number that indicates the cluster. The linear function of the cluster distribution density is:

$$p(X) = \sum_{k=1}^{K} \pi_k f(X|\mu_k, \sum_k) \tag{5}$$

2) E-Step is evaluating the log-likelihood results using the parameter $\mu_k, \sigma_k,$ and π_k. Suppose the cluster C_k represented by a Gaussian distribution (μ_k, σ_k), The probability of X_i in cluster C_k can be calculated by:

$$z_{ik}/\rho(C_k|x_i) = \frac{\rho(x_i|C_k)\rho(C_k)}{\rho(x_i)} \tag{6}$$

Likelihood value:

$$(C_k|x_i) = \frac{1}{\sqrt{2\pi_k}\sigma} e^{\frac{-(x_i-\mu_k)^2}{2\sigma^2}} \tag{7}$$

$$\rho(x_i) = \sum_k \rho(x_i|C_k)\rho(C_k) \tag{8}$$

3) M-Step is changing the value of $\mu_k, \sigma_k,$ and $\rho(C_k)$ with the following calculation:

$$\mu_k = \frac{\sum_i (C_k|x_i)x_i}{\sum_i \rho(C_k|x_i)} \tag{9}$$

$$\sigma_k = \frac{\sum_i (C_k|x_i)(x_i - \mu_k)^2}{\sum_i \rho(C_k|x_i)} \tag{10}$$

$$\pi_k = \frac{\sum_i (C_k|x_i)}{n} \tag{11}$$

4) Perform steps 2 and 3 again until the convergence criteria are met. Therefore, set a certain threshold value for the change of mean and variance in successive iterations, so that the cluster members can be clustered by the Maximum a Posteriori (MAP) method.

$$MAP\{\hat{z}_{ik}\} = \begin{cases} 1 & if \max\{\hat{z}_{ik}\} \epsilon\ k \\ 0 & other \end{cases} \tag{12}$$

Selection of the best model in analysis using the Gaussian Mixture Model (GMM) method is

based on the general approach, namely the Bayes Information Criterion (BIC) value for the model of the parameters and the number of clusters formed [16][24][25].

$$BIC_k = 2\log \mathrm{P}\left(\mathrm{y}\middle|\hat{\theta}_k, M_k\right) - V_k \log(n) \approx 2\log\ \mathrm{P}(\mathrm{y}|M_k) \quad (13)$$

$\mathrm{P}(\mathrm{y}|M_k)$: integration of likelihoods for M_k model.

$\mathrm{P}\left(\mathrm{y}\middle|\hat{\theta}_k, M_k\right)$: integrated the maximum mixed likelihood for M_k model.

V_k : number of independent parameters estimated in the M_k model.

The best model and number of clusters are determined based on the highest BIC value.

3. RESULTS

Clustering analysis with sub-districts as observations on the research data, which included the prevalence of diarrhea in children under five, percentage of households with clean and healthy living behaviours, percentage of households with healthy latrine facilities, population density per-hectare, and the baby < 6 months old who is exclusively breastfed d, was conducted using R software. The mapping of sub-districts in Bandung City based on the characteristics of each variable can be seen in figure 2, figure 3, figure 4, figure 5, and figure 6.

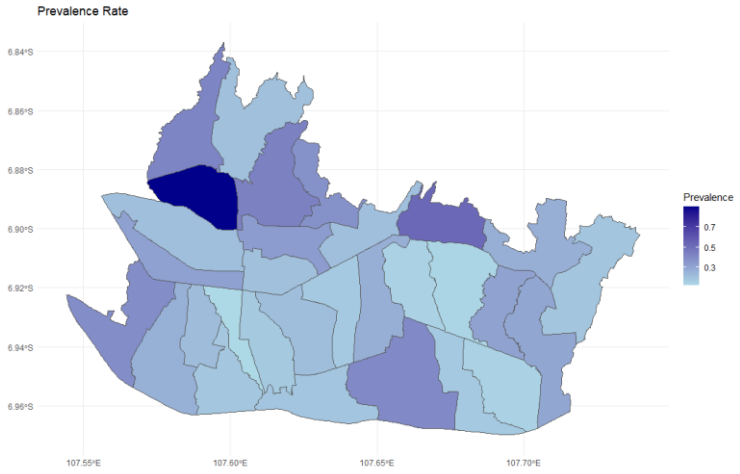

Figure 2. Mapping of Sub-districts based on Prevalence Rate of Diarrhea in Bandung City.

The sub-district with the highest prevalence of diarrhea in children under five is Cinambo at 0.895.

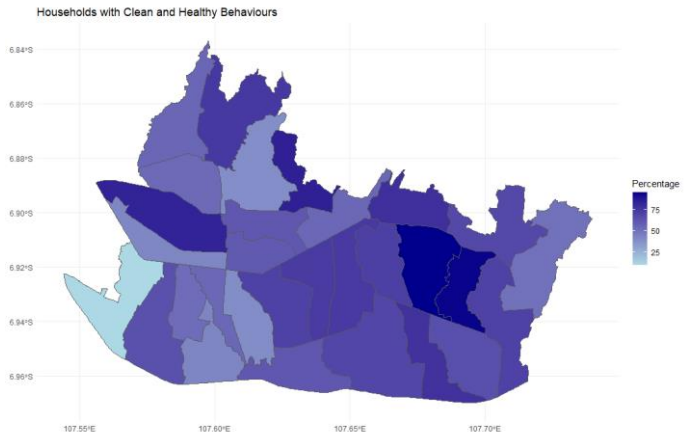

Figure 3. Mapping of Sub-districts based on Percentage of Households with Clean and Healthy Living Behaviours in Bandung City.

25 sub-districts in Bandung City have percentage of households with clean and healthy behaviours above 50%, while 5 sub-districts with low percentage of households with clean and healthy behaviours are Cibeunying Kidul, Cidadap, Cicendo, Bojongloa Kidul, and Bandung Wetan.

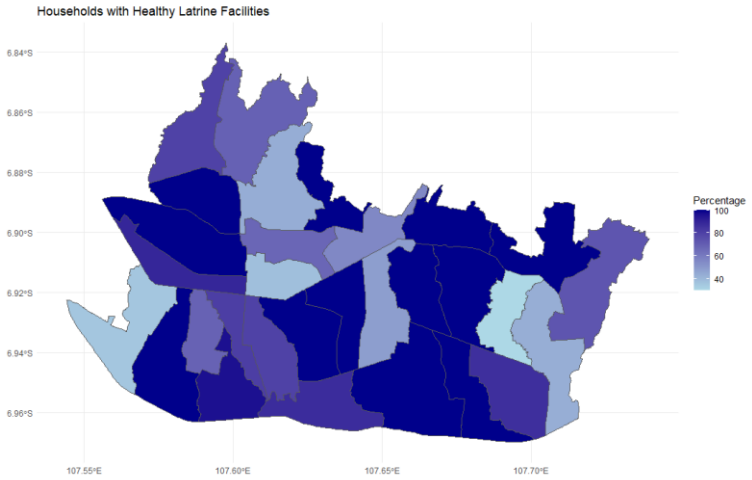

Figure 4. Mapping of Sub-districts based on Percentage of Households with Healthy Latrine Facilities in Bandung City.

12 sub-districts in Bandung City have a percentage of households with healthy latrine facilities 100% including Buah Batu, Panyileukan, Antapani, Mandalajati, Gedebage, Rancasari, Lengkong, Arcamanik, Sukasari, Ujung Berung, Cibiru, and Cinambo.

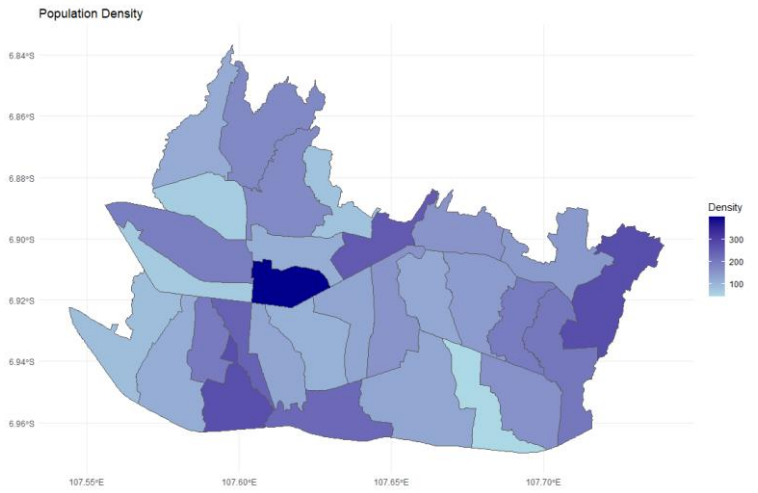

Figure 5. Mapping of Sub-districts based on Population Per-hectare in Bandung City. The highest population density at 399 per-hectare is Sub-district of Bojong Kaler.

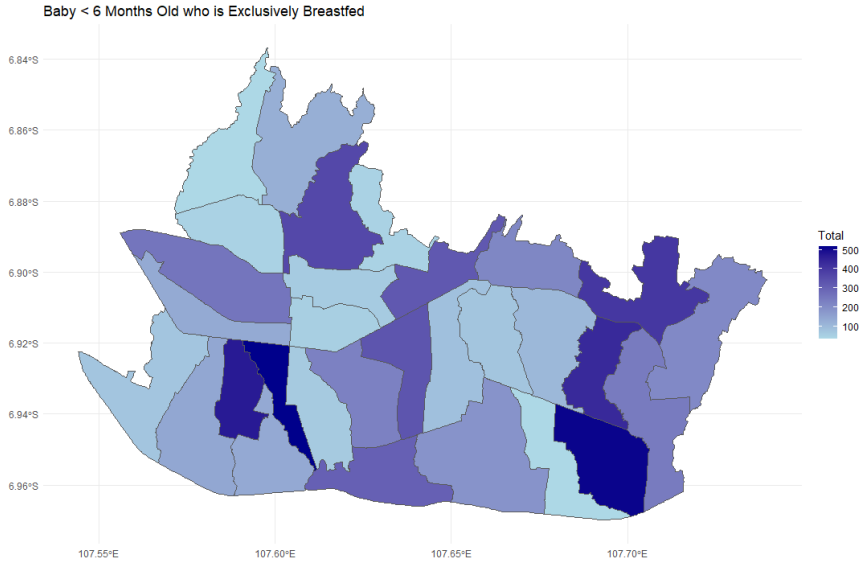

Figure 6. Mapping of Sub-districts based on Number of Baby less than 6 Months Old who is Exclusively Breastfed in Bandung City.

There are 5 sub-districts in the city of Bandung with a high number of babies less than 6 months old who is exclusively breastfed above 400 babies, namely Andir, Coblong, Bandung Kulon, Sukajadi, and Ujung Berung sub-districts.

3.1. Multicollinearity Test

Testing the multicollinearity between variables is carried out as an initial stage in determining

variable selection. If there is multicollinearity in the variables used, it is necessary to select variables with related methods, such as Principle Component Analysis (PCA). The results of multicollinearity testing using the Variance Inflation Factor (VIF) can be viewed in table 2.

Table 2. VIF value

Variables	VIF
Prevalence of diarrhea in children under five	1.268
Households with Clean and Healthy Living Behaviours	1.129
Households with Healthy Latrine Facilities	1.368
Population density	1.542
baby <6 months old who is exclusively breastfed	1.150

Based on table 2, not found the value of VIF more than 10. So it can be concluded that there is no multicollinearity in the variables and all variables will be used in cluster analysis.

3.2. Identification of BIC Value

The identification of the BIC value is done to determine the best model and the number of clusters formed. Based on the analysis results, a comparison chart of the BIC values of various models was obtained. The Ellipsoidal, Equal Volume and Shape (EEV) model has the highest value BIC based on figure 7.

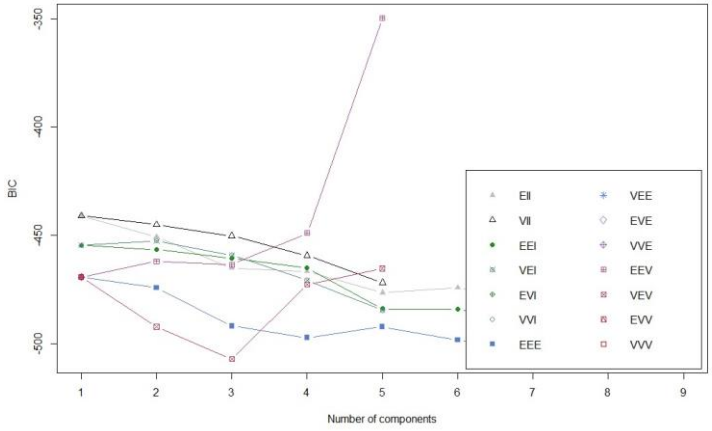

Figure 7. BIC Value of GMM results.

The identification results of the Gaussian Mixture Model EEV based on the Expectation-Maximization (EM algorithm show a model with five components which can be viewed in Table 3.

Table 3. Five component of EEV model

Loglikehood	n	df	BIC	ICL
-31.9397	30	84	-349.58	-349.5899

Bayes Information Criterion (BIC) and Integrated Completed Likelihood (ICL) are metrics used in Gaussian Mixture Model (GMM) analysis to identify the most appropriate number of Gaussian components. In this study, the BIC value is used as a metric in determining the optimal number of components. The BIC value of -349.58 is the highest value in the EEV model with 5 Gaussian components formed, meaning that sub-districts in Bandung City can be grouped into 5 clusters based on predetermined variables.

3.3. Clustering

The grouping of sub-district into 5 clusters based on the EEV model shows that cluster 1 consists of 6 sub-district, cluster 2 consists of 9 sub-district, cluster 3 consists of 5 sub-district, cluster 4 consists of 4 sub-district, and cluster 5 consists of 6 sub-district with the characteristics of each cluster can be viewed in table 4.

Table 4. Means of cluster

Variabel	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Prevalence of Diarrhea in Children Under Five	-0.345	-0.051	-0.547	-0.057	0.915
Households with Clean and Healthy Living Behaviours	-0.238	0.790	-0.508	0.849	-1.089
Households with Healthy Latrine Facilities	0.037	0.884	-0.380	-1.328	-0.158
Population Density	0.149	-0.547	1.643	0.266	-0.877
Baby < 6 Months Old who is Exclusively Breastfed	1.476	-0.420	-0.001	0.098	-0.905

The values in table 4 show the average value that represents the centre of the component in the data space. The average of the variables in each Gaussian component is different from each other, indicating that the distribution of sub-districts in each cluster is based on the extent to which sub-district characteristics are close to the centre of a particular cluster.

The grouping of 30 sub-districts into 5 clusters can be viewed in table 5, and the cluster mapping can be viewed in figure 8.

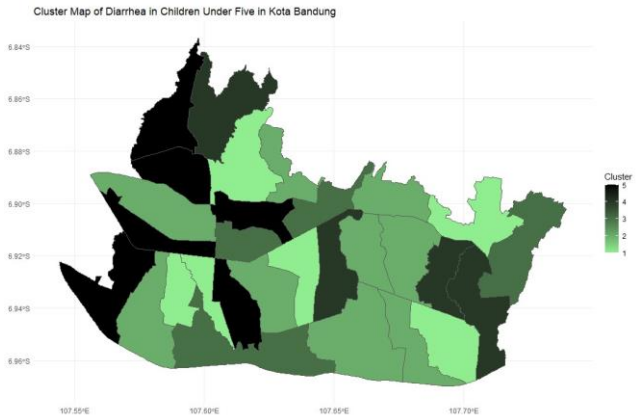

Figure 9. Cluster map of diarrhea in children under five in Kota Bandung.

Based on Figure 9, the different colours of the regions indicate the different clusters. Areas with black colour are areas characterised by a high prevalence of diarrhea.

Table 5. Clustering and characteristics of each cluster

Cluster	Characteristics	Sub-districts
1	Components in cluster 1 tend to be characterized by a high number of baby less than 6 months old who is exclusively breastfed.	Andir, Bandung Kulon, Bojongloa Kidul, Coblong, Rancasari, and Ujung Berung.
2	Components in cluster 2 tend to be characterized by a high percentage of households clean and healthy living behaviours, and also healthy latrine facilities.	Antapani, Arcamanik, Buah Batu, Cibiru, Gedebage, Lengkong, Mandalajati, Panyileukan, and Sukasari.
3	Components in cluster 3 tend to be characterized by high population density per-hectare.	Astana Anyar, Batununggal, Bojong Kaler, Cibeunying Kidul, and Kiaracondong,
4	Components in cluster 4 tend to be characterized by a high percentage of households clean and healthy living behaviours and a high per-hectare population density.	Babakan Ciparay, Cibeunying Kaler, Regol, and Sukajadi
5	Components in cluster 5 tend to be characterized by a high prevalence of diarrhea among under-fives, as well as a low percentage of households clean and healthy living behaviours and number of baby less than 6 months old who is exclusively breastfed.	Bandung Kidul, Bandung Wetan, Cicendo, Cidadap, Cinambo, and Sumur Bandung.

4. CONCLUSION

Clustering was carried out in 30 sub-districts based on 5 variables studied including the prevalence of diarrhea in children under five, the percentage of households with clean and healthy living behaviors, the percentage of households with healthy latrine facilities, the population per-hectare, and the number of baby less than 6 months old who are exclusively breastfed. Based on clustering using the Gaussian Mixture Model (GMM) method with the best model being Ellipsoidal, Equal Volume and Shape (EEV) as many as 5 components, it is concluded that cluster 5 is a group of areas that have a high prevalence of diarrhea in children under five in 2022 so that the need for intervention and efforts of the Bandung City government in preventing higher cases of diarrhea in children under five in six sub-districts including Bandung Kidul, Bandung Wetan, Cicendo, Cidadap, Cinambo, and Sumur Bandung. Cluster 2 consists of sub-districts with good sanitation, including Antapani, Arcamanik, Buah Batu, Cibiru, Gedebage, Lengkong, Mandalajati, Panyileukan, and Sukasari.

It is important to educate the sub-districts included in cluster 5 about better hygiene and sanitation facilities to prevent an increase in under-five diarrhea cases to achieve the Sustainable Development Goals (SDGs) and improve the level of children under five health in Bandung City.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] N.J. CaJacob, M.B. Cohen, Update on diarrhea, Pediatr. Rev. 37 (2016), 313-322. https://doi.org/10.1542/pir.2015-0099.

[2] L.R. Schiller, D.S. Pardi, J.H. Sellin, Chronic diarrhea: diagnosis and management, Clinic. Gastroenterol. Hepatol. 15 (2017), 182-193.e3. https://doi.org/10.1016/j.cgh.2016.07.028.

[3] H. Arifin, W. Rakhmawati, Y. Kurniawati, et al. Prevalence and determinants of diarrhea among under-five children in five Southeast Asian countries: Evidence from the demographic health survey, J. Pediatric Nursing. 66 (2022), e37-e45. https://doi.org/10.1016/j.pedn.2022.06.005.

[4] E. Yunitasari, R. Pradanie, H. Arifin, et al. Determinants of stunting prevention among mothers with children aged 6–24 months, Open Access Maced. J. Med. Sci. 9 (2021), 378-384. https://doi.org/10.3889/oamjms.2021.6106.

[5] M.M. Levine, D. Nasrin, S. Acácio, et al. Diarrhoeal disease and subsequent risk of death in infants and children residing in low-income and middle-income countries: analysis of the GEMS case-control study and 12-month GEMS-1A follow-on study, Lancet Glob. Health 8 (2020), e204-e214. https://doi.org/10.1016/S2214-109X(19)30541-8.

[6] K. Takanashi, Y. Chonan, D.T. Quyen, et al. Survei praktik kebersihan makanan di rumah dan diare masa kanak-kanak di Hanoi, Vietnam, J. Kesehatan Popul. Nutr. 27 (2009), 602-611.

[7] M.B. Karo, Perilaku hidup bersih dan sehat (PHBS) strategi pencegahan penyebaran Virus Covid-19, in: Prosiding seminar nasional hardiknas, vol. 1, pp. 1-4, 2020.

[8] N. Rohmah, F. Syahrul, Hubungan kebiasaan cuci tangan dan penggunaan jamban sehat dengan kejadian diare balita, J. Berkala Epidemiol. 5 (2017), 95-106.

[9] S. Tanuwidjaya, Konsep umum tumbuh dan kembang, in: Tumbuh kembang anak dan remaja. Edisi ke-1, h. 1-12, Jakarta: Sagung Seto, 2002.

[10] L. Novita, D.A. Gurnida, H. Garna, Perbandingan fungsi kognitif bayi usia 6 bulan yang mendapat dan yang tidak mendapat ASI eksklusif, Sari Pediatri, 9 (2016), 429-434.

[11] D.H. Vu, K.M. Muttaqi, A.P. Agalgaonkar, A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables, Appl. Energy. 140 (2015), 385-394. https://doi.org/10.1016/j.apenergy.2014.12.011.

[12] R.S. Gómez, A.R. Sánchez, C.G. García, et al. The VIF and MSE in raise regression, Mathematics. 8 (2020), 605. https://doi.org/10.3390/math8040605.

[13] Z. Wahidah, D.T. Utari, Comparison of k-means and gaussian mixture model in profiling areas by poverty indicators, BAREKENG: J. Ilmu Mat. Terapan. 17 (2023), 0717-0726.

[14] L. Scrucca, Identifying connected components in Gaussian finite mixture models for clustering, Comput. Stat. Data Anal. 93 (2016), 5-17. https://doi.org/10.1016/j.csda.2015.01.006.

[15] S. Belciug, D.G. Iliescu, Deep learning and Gaussian Mixture Modelling clustering mix. A new approach for fetal morphology view plane differentiation, J. Biomed. Inform. 143 (2023), 104402. https://doi.org/10.1016/j.jbi.2023.104402.

[16] N. Shen, B. González, Bayesian information criterion for linear mixed-effects models, preprint, (2021). http://arxiv.org/abs/2104.14725.

[17] C.E. Troeger, I.A. Khalil, B.F. Blacker, et al. Quantifying risks and interventions that have affected the burden of diarrhoea among children younger than 5 years: an analysis of the Global Burden of Disease Study 2017, Lancet Infect. Dis. 20 (2020), 37-59. https://doi.org/10.1016/s1473-3099(19)30401-3.

[18] N. Thapar, I.R. Sanderson, Diarrhoea in children: an interface between developing and developed countries, The Lancet. 363 (2004), 641-653. https://doi.org/10.1016/s0140-6736(04)15599-2.

[19] V. Diwan, Y.D. Sabde, E. Byström, et al. Treatment of pediatric diarrhea: a simulated client study at private pharmacies of Ujjain, Madhya Pradesh, India, J. Infect. Develop. Countries. 9 (2015), 505-511. https://doi.org/10.3855/jidc.5694.

[20] G. Mengistu, K. Gietnet, F. Amare, et al. Self-reported and actual involvement of community pharmacy professionals in the management of childhood diarrhea: a cross-sectional and simulated patient study at two towns of Eastern Ethiopia, Clin. Med. Insights Pediatr. 13 (2019), 117955651985538. https://doi.org/10.1177/1179556519855380.

[21] V. Melnykov, R. Maitra, Finite mixture models and model-based clustering, Stat. Surv. 4 (2010), 80-116. https://doi.org/10.1214/09-ss053.

[22] C. Fraley, A.E. Raftery, Model-based clustering, discriminant analysis, and density estimation, J. Amer. Stat. Assoc. 97 (2002), 611-631. https://doi.org/10.1198/016214502760047131.

[23] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. R. Stat. Soc.: Ser. B (Methodol.) 39 (1977), 1-22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.

[24] G. Schwarz, Estimating the dimension of a model, Ann. Stat. 6 (1978), 461-464. https://www.jstor.org/stable/2958889.

[25] R.A. Stine, Model selection using information theory and the MDL principle, Sociol. Methods Res. 33 (2004), 230-260. https://doi.org/10.1177/0049124103262064.

[26] P.N. Tan, M. Steinbach, V. Kumar, Data mining cluster analysis: basic concepts and algorithms, in: Introduction to Data Mining, pp. 487-533, (2013).

[27] A. Athifaturrofifah, R. Goejantoro, D. Yuniarti, Perbandingan Pengelompokan K-Means dan K-Medoids Pada Data Potensi Kebakaran Hutan/Lahan Berdasarkan Persebaran Titik Panas, EKSPONENSIAL, 10 (2020), 143-152.

[28] A.P. Kusuma, D.M. Sukendra, Analisis spasial kejadian demam berdarah dengue berdasarkan kepadatan penduduk, Unnes J. Public Health, 5 (2016), 48-56.

[29] S. Rini, Rohayati, Kejadian Diare pada Bayi dengan Pemberian ASI, J. Nursing, 11 (2015), 153-156.

[30] N.M.E. Wardani, K.A. Witarini, P.J. Putra, et al. Pengaruh Pemberian ASI Eksklusif Terhadap Kejadian Diare pada Anak Usia 1-3 Tahun, J. Medika Udayana, 11 (2022), 12-17.

[31] L.B. Wadu, A.F. Gultom, F. Pantus, Penyediaan Air Bersih Dan Sanitasi: Bentuk Keterlibatan Masyarakat Dalam Pembangunan Berkelanjutan, J. Pendidikan Kewarganegaraan, 10 (2020), 80-88.