https://doi.org/10.20473/mgi.v17i1SP.76–82

REGIONAL MAPPING IN BANGKALAN DISTRICT BASED ON POTENTIAL INDICATORS OF TOTAL STUNTING USING K-MODE CLUSTER ALGORITHM

Adma Novita Sari[1], Dzuria Hilma Qurotu'ain[1], Faradilla Harianto[1], Sa'idah Zahrotul Jannah[1*]
[1]Faculty of Science and Technology, Universitas Airlangga
*E-mail: s.zahrotul.jannah@fst.unair.ac.id

ABSTRACT

Indonesia has the highest number of stunted children in Southeast Asia. Indonesia's stunting prevalence rate is also still higher than the threshold set by WHO (20%). This issue has become a particular concern of the government. Hence, by 2024, the government targets the prevalence of stunting to decrease to 14% under Presidential Regulation no. 72 of 2021. One of the highest stunting cases in Indonesia is Bangkalan Regency, which has a prevalence rate of 38.9%, exceeding the maximum stunting target limit in Indonesia. Therefore, this study aimed to map the area in Bangkalan Regency based on the factors that influence stunting cases in toddlers. This study used data from the National Population and Family Planning Agency (BKKBN) survey in 2021. The results show that several sub-districts in Bangkalan still have poor water availability, sanitation, environmental hygiene, and housing welfare. Mapping the area can help the government provide the right solution for the area's problems.

Keywords*: Stunting, Regional Mapping, Clustering, K-Modes*

INTRODUCTION

Stunting is a condition where a toddler experiences failure to thrive, which can be caused by various things, such as malnutrition, lack of sanitation, and external factors, such as early marriage by the mother. Malnutrition can arise due to the community's condition, which is unaware of its impact on the future of Indonesia's young generation (Winasis, 2018). Stunting is based on several different groupings. Generally, the indicators used are height and weight. (Fuada, 2017). Many other things can influence stunting. Socio-economic conditions and the existence of sanitation that meets hygiene standards significantly affect the chances of stunting. (Pratiwi et al., 2020). Field evidence suggests that having sanitary conditions that do not meet national health requirements can increase the risk of disease. Infection is the most likely cause of disease. This disorder can prevent nutrients from being absorbed by the baby's digestive system, so the baby gains less weight. Stunting can occur if this condition is not treated for a long time. (Directorate of Public Health and Nutrition and Secretary for the Acceleration of Nutrition Improvement - Bappenas, 2018).

Currently, stunting is a significant health problem in Indonesia which is included in one of the discussions of the second point of the Sustainable Developments Goals. Stunting does not only impact the growth of children under five but also the risk of death, obesity, chronic diseases, and even mental health and intellectual intelligence. It affects human resources, which will play a vital role in Indonesia's development in the future. If the human resources of this country are still in poor condition, it is impossible to compete with other countries in addressing global concerns. Therefore, stunting must be taken seriously as a problem. (Teja, 2019).

Many studies conducted research on how to reducing stunting, one of them is evaluating the health and nutrition programs to recognize the top programs for reducing stunting in developing countries. It showed that the implemented programs for reducing stunting are counselling, growth monitoring, immunization, water, sanitation and hygiene, and social safety nets. Government commitment, multi-sectoral collaboration, community engagement, and wider program coverage can help those programs to perform well (Hossain, et al., 2017)

 Received 18-11-2022, Accepted 13-12-2022, Published online 15-12-2022.

In Southeast Asia, Indonesia has the second-highest child stunting rate and ranks sixth globally. With a prevalence of 20%, the stunting rate recorded is higher than the WHO's upper limit. On the other hand, Indonesia experienced a 24.4% increase in stunting cases in 2021. This figure still needs to catch up to the government's target of accelerating stunting reduction with a prevalence of 14% in 2024 as outlined in the Presidential Regulation (Perpres) No. 72 the Year 2021. (Teja, 2022).

One of the areas in East Java Province with the highest stunting rate is Bangkalan Regency. Reported from the Bangkalan Regency PPKBP3A Office, Amina Rahmawati stated in 2021 that the stunting prevalence rate in Bangkalan Regency based on the results of the SSGI (Indonesian Nutritional Status Study) data reached a value of 38.9%. This figure is high enough to exceed Indonesia's maximum stunting target limit

Covid-19 pandemic that emerged in 2020 caused all lines of life to be hampered, including the health sector. Stunting cases are estimated to increase due to several factors. During the Covid-19 pandemic, several cases attacked pregnant women. During childbirth, there is a possibility that the mother is suffering from Covid-19, so the mother cannot breastfeed, which will determine the baby's life in the future. In addition, during the Covid-19 pandemic, pregnant women are hampered from being able to carry out regular consultations with doctors, which also affects the condition of the baby they are carrying. Moreover, the case of pregnant women who died from Covid-19 during childbirth also affected the condition of the baby.

This research aims to map which areas have the highest stunting potential, considering that Bangkalan Regency is the area in East Java with the most stunting cases. This research is expected to minimize stunting cases by reviewing and providing follow-up to areas with great potential for stunting. This effort is also to support sustainable development or SDGs and improve the quality of Indonesian Human Resources in the future.

Regional mapping in Bangkalan Regency will help government to prioritize the implementation of programs for reducing stunting. Nambiar, et.al (2022) showed that social demographic indicators for spatial clustering played important role to define priority programs in specific area. Therefore, this research analyzes the characteristics of each cluster and give recommendations based on its conditions.

METHOD

The source of data used in this research design is secondary data obtained from the National Population and Family Planning Agency of East Java Province in 2021. The data obtained contains factors that influence stunting in East Java. This research focuses on data from Bangkalan Regency, the area in East Java with the greatest risk of stunting. The variables are based on indicators of stunting, such as marital age, family welfare, nutrition, savings ownership, type of housing, water source, and the existence of sanitation that meets national standards based on drinking water sources and sanitation ownership.

Data analysis used two types of statistical analysis, namely factor analysis, and continued with cluster analysis. Most variables are categorical data, so the K-Modes algorithm is suitable for the analysis.

Factor Analysis

Factor analysis is a technique to find out how many independent variables relate to each other. Later, these variables will be simplified based on this research. This analysis aims to establish the relationship between variables using a correlation test and create a group of variables to replace the others.

The assumptions used in factor analysis are:

a. The magnitude of the correlation coefficient between variables must be greater than 0.05
b. The magnitude of the partial correlation or between two variables (Anti Image Correlation) must be small
c. The correlation between variables (Bartlett Test of Sphericity or Measure Sampling Adequacy) must be significant as a condition for further analysis.
d. Factor analysis by extracting one or more factors by providing interpretation and description of the variables.

e. Validate the output and compare it to determine whether the differences formed are valid by using the Confirmatory Factor Analysis (CFA) method. (Yuliandi, 2013)

Cluster Analysis

Cluster analysis is a technique or test method used to group objects based on their characteristics. This analysis attempts to group objects with the most similar characteristics so that they will be included in the same cluster. The requirement for this analysis is that the data must have high homogeneity. In cluster analysis, a set of variables is known that describes the characteristics possessed.

A fundamental difference separates cluster analysis and factor analysis. While factor analysis concentrates on grouping variables, cluster analysis concentrates on grouping objects. When performing cluster analysis, specific outputs are generated, and cluster members depend on various factors related to the procedure. Modifying one or more of these factors can generate alternative results. When a variable has an impact on cluster results, a change occurs.

Because the data used is categorical, the K-Modes methodology is used in this cluster analysis method. This method is the result of changing the standard k-means algorithm into a categorical data cluster by substituting a distance function unsuitable for the Euclidean distance function. The K-Modes function, which is called by installing the Python software, is used for data processing with cluster analysis.

The stages of clustering using the K-Modes algorithm are as follows Huang, 2008):

a. Determines k clusters and performs cluster initialization randomly.
b. Allocating data objects to clusters based on the closest distance based on a simple dissimilarity measure. Then do regular updates.
c. Re-check each object's dissimilarity value against the mode to avoid miss-clustering.
d. Repeat the third step until no data objects change the cluster.

RESULT AND DISCUSSION

Factor Analysis

Table 1. Result of KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		**.747**
Bartlett's Test of Sphericity	Approx. Chi-Square	59684.275
	Df	21
	Sig.	.000

The Kaiser Meyer Olkin Measure of Sampling (KMO) and Bartlett's Test score is a comparison index of the distance between the correlation coefficient and the partial correlation coefficient is 0.747 with a significance of 0.000. Because the number is already above 0.5 and the significant number is below 0.05 (α = 5%), the existing variables and samples can be analyzed further.

Table 2. Anti image Correlation

Anti-image Matrices

		usia_kawin	sumber_penghasilan	gizi	tabungan	alas_rumah	sumber_airminum	sanitasi
Anti-image Covariance	usia_kawin	.920	.049	7.756E-5	.023	-.182	-.026	.000
	sumber_penghasilan	.049	.490	-.295	-.134	-.054	-.027	-.016
	gizi	7.756E-5	-.295	.519	-.080	.012	-.039	-.051
	tabungan	.023	-.134	-.080	.731	-.111	-.004	-.062
	alas_rumah	-.182	-.054	.012	-.111	.641	-.187	-.195
	sumber_airminum	-.026	-.027	-.039	-.004	-.187	.753	-.155
	sanitasi	.000	-.016	-.051	-.062	-.195	-.155	.720
Anti-image Correlation	usia_kawin	.600[a]	.074	.000	.028	-.237	-.031	.000
	sumber_penghasilan	.074	.688[a]	-.584	-.224	-.096	-.044	-.026
	gizi	.000	-.584	.692[a]	-.130	.021	-.063	-.084
	tabungan	.028	-.224	-.130	.858[a]	-.163	-.005	-.086
	alas_rumah	-.237	-.096	.021	-.163	.745[a]	-.269	-.287
	sumber_airminum	-.031	-.044	-.063	-.005	-.269	.811[a]	-.211
	sanitasi	.000	-.026	-.084	-.086	-.287	-.211	.811[a]

a. Measures of Sampling Adequacy(MSA)

Table 2 shows the significant correlation number that indicates the Bartlett Test of Sphericity or Measure Sampling Adequacy (MSA) number. The explanation of the table is as follows:

1. The MSA figure for the variable of marital age is 0.60, for the source of income variable is 0.688, and so on.
2. It can be shown that each variable has MSA value that more than 0.5. Thus, a deeper analysis of these variables is possible.

Table 3. Result of Extraction Method

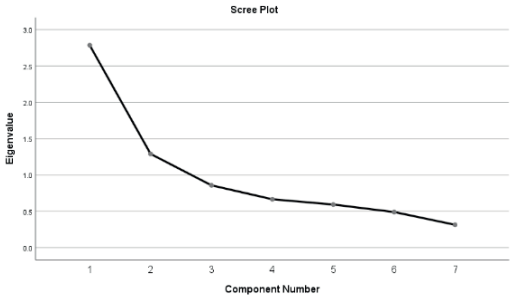
Communalities		
	Initial	**Extraction**
usia_kawin	1.000	.466
sumber_penghasilan	1.000	.750
Gizi	1.000	.711
Tabungan	1.000	.494
alas_rumah	1.000	.657
sumber_airminum	1.000	.491
Sanitasi	1.000	.506
Extraction Method: Principal Component Analysis.		

1. The value is 0.466 for variable 1, which indicates the age of marriage. It shows that the formed factors can contribute about 46.6% of the variation in the variable age at marriage.
2. The value is 0.750 for variable 2, or source of income. It shows that the resulting component can account for about 75% of the variance of the income source variable.
3. The result for variable 3, namely nutrition, is 0.711. It indicates that the resulting component can account for about 71.1% of the variance of the nutrition variable
4. The result for variable 4, namely savings, is 0.494. It shows that the resulting component can account for about 49.4% of the variance of the savings variable.
5. The result for variable 5, namely the base of the house, is 0.657. This shows that the resulting component can account for about 65.7% of the variance of the house base variable.
6. The result for variable 6, namely the source of drinking water, is 0.491. This shows that the resulting component can contribute about 49.1% of the variance of the drinking water source variable.
7. The result for variable 7, namely sanitation, is 0.506. This shows that the resulting component can account for about 50.6% of the variance of the sanitation variable.

Table 4. Sum of Variance Explained

Total Variance Explained									
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.786	39.799	39.799	2.786	39.799	39.799	2.204	31.490	31.490
2	1.290	18.430	58.229	1.290	18.430	58.229	1.872	26.739	58.229
3	.858	12.259	70.488						
4	.666	9.518	80.006						
5	.594	8.484	88.491						
6	.491	7.009	95.499						
7	.315	4.501	100.000						
Extraction Method: Principal Component Analysis.									

1. Five variables (components), X1 (age of marriage), X2 (source of income), X3, (nutrition), X4, (savings), X5, (house mat), X6, and X7 (drinking water supply—included in factor analysis (sanitation).
2. Pay attention to the "Initial Eigenvalues" column, where set the value to 1 using SPSS (one). The variance that can be explained by a factor of one is $\frac{2.786}{7} \times 100\%$, or 39.799%. $\frac{1.290}{7} \times 100\%$, or a factor of 2, equals 18.430%. So that the 39.799% + 18.430% = 58.22% variable can be explained by the addition of the two elements.



Graph 1. Chart of Scree Plot

Graph 1 shows that the orientation of the line is dropped dramatically from 1 factor to 2 factors (line from Component Number = 1 to 2 axes). The trend of the line continues to decrease as the factor increases to 3, 4, 5, and so on. This indicates that for summarizing the five variables, two components are the most effective.

Table 5. Matrix Components Based on the Results of Factor Analysis

Component Matrix[a]		
	Component	
	1	**2**
usia_kawin	.180	.659
sumber_penghasilan	.740	-.450
Gizi	.715	-.447
tabungan	.659	-.245
alas_rumah	.685	.434
sumber_airminum	.606	.352
sanitasi	.652	.284
Extraction Method: Principal Component Analysis.		
a. 2 components extracted.		

The Component Matrix shows the distribution of the seven variables on the two formed factors. While the numbers in the table are factor loadings

which show the magnitude of the correlation between a variable in factors 1 and 2.

Table 6. Rotated Component Matrix

Rotated Component Matrix[a]		
	Component	
	1	2
usia_kawin	-.270	.627
sumber_penghasilan	.859	.109
Gizi	.838	.096
Tabungan	.668	.219
alas_rumah	.265	.766
sumber_airminum	.254	.653
Sanitasi	.332	.628

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.[a]
a. Rotation converged in 3 iterations.

From table 6, it can be seen that the correlation between the factors that are most correlated with factor 1 is the source of income: 0.859, while the most correlated with factor 2 is the base of the house: 0.766. Then it can be concluded that the members of each factor:

Factor 1: source of income, nutrition, savings
Factor 2: age of marriage, source of drinking water, bedding, ownership of sanitation

Table 7. Component Transformation Matrix

Component Transformation Matrix		
Component	1	2
1	.782	.624
2	-.624	.782

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

The table above shows that in component 1, the correlation value is 0.782 > 0.5, and in component 2: 0.782 > 0.5. Because all components > 0.5, the two factors formed can be said to be right in summarizing the seven existing variables.

Based on the analysis of the factors above, it can be concluded that:

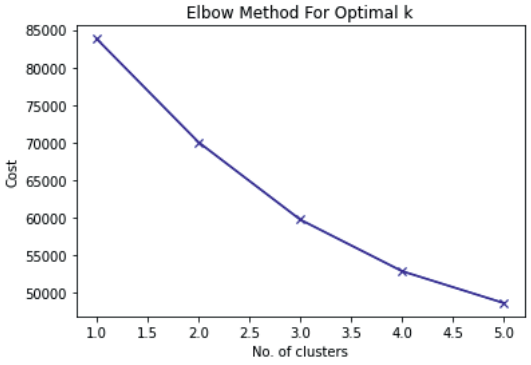
1. Of the seven variables studied, it turned out that there were no variables that were not worthy of further analysis. So that all variables are included in the following factor analysis.
2. Of the seven variables studied, the factoring process can be reduced to only two factors

Factors formed:

Factor 1: source of income, nutrition, savings
Factor 2: age of marriage, source of drinking water, bedding, sanitation

Cluster Analysis

Cluster analysis is used to see the characteristics of the data that has been obtained. This analysis can be carried out on various values of k (k=2,3,4 and 5). In determining the best alternative k based on the rules of the elbow method. The following is the output result of the Elbow Method, which is shown in graph 2:



Graph 2. Elbow Method

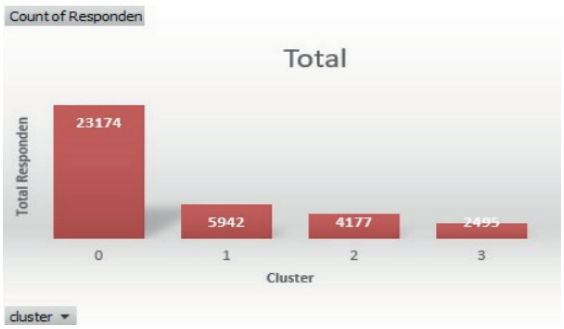
Graph 2 shows that the best alternative value for k is k=4 since there is a sloping shape like an elbow at that value. Therefore, a cluster test with many k = 4. The results of the cluster processing are shown in table 8.

Table 8. Result of Cluster Analysis using K-Modes Method

usia_kawin	sumber penghasilan	gizi	Tabungan	Alas rumah	Sumber air minum	Sanitasi	cluster	
0	tdk ideal	tidak ada	tercukupi	ada	1	1	1	0
1	tdk ideal	tidak ada	tercukupi	ada	1	2	1	0
2	tdk ideal	tidak ada	tercukupi	ada	1	2	1	0
3	tdk ideal	tidak ada	tidak tercukupi	tidak ada	1	2	1	3
4	tdk ideal	tidak ada	tidak tercukupi	ada	1	2	1	0
...	...	...	...	...	...	...	...	...
35783	ideal	ada	tercukupi	tidak ada	2	4	2	3
35784	ideal	ada	tercukupi	ada	2	4	3	0
35785	ideal	ada	tercukupi	ada	2	4	3	0
35786	ideal	ada	tercukupi	ada	2	4	3	0
35787	ideal	ada	tercukupi	tidak ada	5	4	2	0

35788 rows × 8 columns

There are 23174 respondents in cluster 1, 5942 respondents in cluster 2, 4177 respondents in cluster 3, and 2495 respondents in cluster 4.



Graph 3. Respondent Distribution Graph

The cluster can be a reference for grouping sub-district in Bangkalan Regency. The characteristics of each cluster are based on the mode of each variable. Table 9 shows the district cluster and recommendation based on the cluster characteristics.

Table 9. District Clusters and Recommendations for Each Cluster

Cluster	District Cluster member	Cluster Characteristics	Recommendations
CLUSTER 1	Bangkalan, Konang, Arosbaya, Burneh, Galis, Geger, Klampis, Kokop, Konang, Sepulu, Tanah merah, Tanjung Bumi	• Mother's understanding of the ideal marriage is good • Parents' nutritional needs are fulfilled • Availability of water, sanitation, and good environmental hygiene • Families have sources of income • Welfare of housing is good	• Implement stunting-oriented policies • Conducting outreach gradually to the community so that they are not oblivious to various variables that can increase the potential for stunting • Stunting alleviation program • Educate the local community regarding the importance of sanitation and hygiene in the family environment • The aid program is allocated for the improvement of bathing, washing, and latrine facilities
CLUSTER 2	Kamal, Socah	• Mother's understanding of the ideal marriage is good • Parents' nutritional needs are fulfilled • Families have sources of income • Poor water availability, sanitation and environmental hygiene • Welfare of housing is good	• Stunting alleviation program • Educate the local community regarding the importance of sanitation and hygiene in the family environment • • The aid program is allocated for the improvement of bathing, washing and latrine facilities
CLUSTER 3	Tragah, Blega	• Mother's understanding of the ideal marriage is not good • Parents' nutritional needs are fulfilled • Families have sources of income • Availability of water, sanitation, and good environmental hygiene • Poor housing welfare	• Guidance and sharing programs regarding education about marriage that parents should pay attention to • Cultivation and counseling related to the ideal marriage and cleanliness of the residence • The aid program is allocated for the improvement of residential facilities
CLUSTER 4	Modung, Labang, Kwanyar	•Mother's understanding of the ideal marriage is not good • Parents' nutritional needs are fulfilled • Families have no source of income • Availability of water, sanitation, and good environmental hygiene • Welfare of housing is good	• Guidance and sharing programs regarding education about marriage that parents should pay attention to • Cultivation and counseling related to the ideal marriage • Job training programs • • Job creation

CONCLUSION

Cluster analysis using the K-Modes method can be an alternative to analyze stunting cases in Bangkalan Regency. The best alternative number of clusters (k) is k=4. Based on the results, Konang,

Arosbaya, Burneh, Galis, Geger, Klampis, Kokop, Konang, Sepulu, Tanah Merah, and Tanjung Bumi sub-districts are included in Cluster 1 with the characteristics of indicators of stunting potential still good. Kamal and Socah sub-districts are in Cluster 2, with the characteristics of indicators of stunting potential being quite good. However, there are indicators of water and sanitation hygiene that can increase stunting potential. Tragah and Blega sub-districts are in Cluster 3 with indicators that trigger stunting, namely parental understanding regarding stunting, early marriage, and housing welfare. Meanwhile, Modung, Labang, and Kwanyar sub-districts are in Cluster 4 with indicators that trigger stunting, namely parental understanding regarding stunting, early marriage, and income source problems.

Policies and various development programs need to be carried out by various essential elements in the community, both from the government, village administrators, students, and the Bangkalan community itself, to suppress the potential that can trigger stunting in toddlers. Education programs need to be carried out to increase public understanding and concern about the seriousness of the stunting problem in Indonesia. Job training programs and job creation also need to be organized to support sources of income and the basic needs of the community in order to increase community welfare.

REFERENCES

Aryastami, N. K., & Tarigan, I. (2017). Kajian Kebijakan dan Penanggulangan Masalah Gizi Stunting di Indonesia. Buletin Penelitian Kesehatan, 45(4), 233–240. https://doi.org/10.22435/bpk.v45i4.7465.233-240

Direktorat Kesehatan dan Gizi Masyarakat dan Sekretaris Percepatan Perbaikan Gizi - Bappenas. (2018). Pencegahan Stunting dan Pembangunan Sumber Daya Manusia. In Buletin Jendela Data dan Informasi Kesehatan (Vol. 53, Issue 9, pp. 38–43). https://doi.org/10.1017/CBO9781107415324.004

Fuada, N. (2017). STATUS GIZI ANAK BADUTA (BAWAH DUA TAHUN) DI INDONESIA. *Jurnal Litbang Provinsi Jawa Tengah*, *15*(1), 51-63.

Huang, N. E., & Wu, Z. (2008). A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Reviews of geophysics*, *46*(2).

Hossain, M., Choudhury, N., Abdullah, K. A., Mondal, P., Jackson, A. A., Walson, J., & Ahmed, T. (2017). Evidence-based approaches to childhood stunting in low and middle income countries: a systematic review. *British Medical Journal*, 903-909.

Nambiar, A., Agnihotri, S. B., Singh, A., & Arunachalam, D. (2022). Region matters: Mapping the contours of undernourishment among children in Odisha, India. *PLOS ONE*, 1-21.

Pratiwi, W. R., Hamdiyah, & Asnuddin. (2020). Deteksi Dini Masalah Kesehatan Reproduksi Melalui Pos Kesehatan Remaja. Jurnal Inovasi Hasil Pengabdian Masyarakat (JIPEMAS), 3(1), 87–94. https://doi.org/10.33474/jipemas.v3i1.5035

Satriawan, E. (2018). Strategi Nasional Percepatan Pencegahan Stunting 2018-2024. *Jakata: Tim Nasional Percepatan Penanggulangan Kemiskinan (TNP2K)*.

Teja, (2019). Analisis Structural Equation Modeling (SEM) untuk Sampel Kecil dengan Pendekatan Partial Least Square (PLS).

Winasis, N. P. (2019). *Analisis Faktor Kejadian Stunting pada Anak Usia 24-59 Bulan Berbasis Transcultural Nursing di Desa Morombuh Kecamatan Kwanyar Bangkalan* (Doctoral dissertation, Universitas Airlangga).