J-COSINE (Journal of Computer Science and Informatics Engineering)
Vol. 10, No. 1, June 2026 E-ISSN:2541-0806
Accredited Sinta-4 by RISTEKDIKTI Decree No. 79/E/KPT/2023 P-ISSN:2540-8895

# Provincial Segmentation Based on K-Medoids Clustering for Risk Factor Mapping and Targeted Intervention of Stunting in Indonesia

Dina Eka Putri[1]*, Syifa Salsabila Satya Graha[2], Baiq Fitria Rahmiati[3]
[1]Dept. of Statistics, Universitas Mataram
Jl. Majapahit 62, Mataram, Lombok NTB, INDONESIA
[2]Dept. of Mathematics, Universitas Mataram
Jl. Majapahit 62, Mataram, Lombok NTB, INDONESIA
[3]Dept. of Nutrition, Universitas Bumi Gora
Jl. Ismail Marzuki 22, Mataram, Lombok NTB, INDONESIA
*Email:* dina.putri@unram.ac.id, syifbilstgrha423@gmail.com, baiqfitria@universitasbumigora.ac.id
***Corresponding Author***

***Abstract*** **Stunting is a multidimensional, chronic nutritional problem shaped by maternal health, caregiving practices, and environmental conditions. This study segments all 38 Indonesian provinces based on ten stunting risk indicators from the 2023 Indonesian Health Survey (SKI) using K-Medoids Partition Around Medoids (PAM) with Median-MAD robust scaling and Manhattan distance. Four optimal clusters were identified (Average Silhouette Width, ASW = 0.425). Variable contribution analysis using Kruskal-Wallis $\eta^2$ and permutation-based $\Delta$ASW identified maternal Chronic Energy Deficiency (CED; $\eta^2$ = 0.535) and maternal anemia prevalence ($\eta^2$ = 0.435) as the dominant cluster discriminators, followed by ANC $\geqslant$10 visits, LBW, exclusive breastfeeding, and wasting in adult women. Sanitation access structurally separates the Papua highlands cluster. Four distinct provincial profiles emerged: (1) good WASH, moderate ANC and breastfeeding gaps; (2) good WASH with high maternal nutritional burden; (3) high anemia despite good service coverage; and (4) severe WASH deficit with low ANC. Cluster-specific policy recommendations are aligned with Indonesia's existing programmes, i.e. Asta Cita 4, Gerakan Nasional Percepatan Perbaikan Gizi (Gernas PPG), the Thousand Days of Life (HPK) programme, and SDG 2, with targeted enhancements addressing identified gaps.**

***Key words*****: Stunting, Unsupervised Learning, K-Medoids Clustering, ANC Health, WASH**

## I. Introduction

Stunting is a chronic, multidimensional nutritional problem. It begins before pregnancy, continues throughout pregnancy, and continues into early childhood, influenced by parenting practices, maternal nutritional status, and environmental conditions. This condition is characterized by long-term malnutrition caused by inadequate nutritional intake [1]. Based on the 2024 Indonesian Nutritional Status Survey (SSGI), the national prevalence of stunting has decreased to 19.8% [2]. According to World Health Organization (WHO) standards, this prevalence rate is considered medium, at less than 20% [5]. However, the decline has not been evenly distributed across all provinces in Indonesia. This inter-regional variation indicates that the causes of stunting are complex and multifactorial, making understanding the mapping of risk characteristics in each province crucial for designing more targeted interventions.

The causes of stunting can be categorized into three levels: direct, intermediate, and underlying causes. Direct causes include micronutrient intake and infection. Intermediate causes include maternal health status, food access, and morbidity. Meanwhile, underlying causes involve socio-economic conditions, sanitation, and education [4][5]. Various previous studies have shown that strengthening maternal nutrition during pregnancy through increasing Antenatal Care (ANC) visits, iron-folic acid (Fe/IFA) or multi-micronutrient (MMS) supplementation, and nutrition education contribute significantly to reducing anemia and increasing hemoglobin levels [6][7]. The effectiveness of nutrition interventions will be more optimal by combining improvements in Water, Sanitation, and Hygiene (WASH) conditions. A cross-cluster approach that combines strengthening ANC, maternal nutrition services, WASH interventions, and community advocacy and empowerment is expected to have a greater nutritional impact than partial interventions.

Efforts to eradicate stunting align with the global commitment through the Sustainable Development Goals (SDGs) established by the United Nations (UN) in September 2015. The SDGs have 17 global goals and objectives, with 169 targets and 289 indicators, aimed at improving the well-being of society by 2030 [8]. One of these is the second goal (SDG 2), with target 2.2, which focuses on eliminating all forms of malnutrition. This commitment also aligns with Indonesia's Asta Cita 4, which emphasizes improving the nutritional status of mothers and children as part of the national health development priorities.

In the context of data analysis and evidence-based decision-making, the use of the K-Medoids Clustering method is relevant for segmenting provinces in Indonesia based on stunting risk. This method was chosen because it is robust, with outliers present, requiring a complete number of provinces, and is capable of grouping aggregate data between provinces based on shared characteristics. The use of robust scale (Median-MAD) and Manhattan distance in this analysis allows for more stable clustering results compared to the K-Means method. Based on this, this study aims to conduct an unsupervised learning-based segmentation approach for provinces in Indonesia using the K-Medoids method using ten stunting risk indicators [9]. This approach is expected to serve as a basis for formulating more effective, contextual, and appropriate nutrition policies and interventions tailored to the characteristics of each province.

## II. Literature Review

### *A. Related Research*

Clustering methods have been applied to regional stunting segmentation in Indonesia at multiple scales. Pohan et al. (2021) grouped 34 Indonesian provinces by child nutritional status using K-Medoids, yielding two clusters — a majority with better outcomes and a smaller high-burden cluster concentrated in eastern Indonesia [11]. At sub-district level, Cahyani et al. (2024) identified five clusters in Yogyakarta (DIY), differentiated by income, water access, toilet availability, and stunting case counts [10]. Fadhilah et al. (2025) applied Principal Component Regression across seven indicators, explaining 63.3% of national stunting variation through two principal components [9]. These studies confirm that multi-indicator, province-level segmentation is both methodologically feasible and policy-relevant.

In the nutrition domain, Billah et al. (2022) in Bangladesh found that despite 80% IFA receipt, only 8% consumed the recommended ≥180 tablets; ANC attendance, nutrition education, and free access were independently associated with adequate coverage [6]. A meta-analysis by Engidaw et al. (2024) confirmed that nutrition education via ANC improved IFA adherence (OR ≈ 2.80), raised haemoglobin by 0.88 g/dL, and reduced anemia risk by approximately 34% [7]. On the WASH side, Bekele et al. (2020) found that standalone WASH interventions without nutritional components had limited effect on anemia [12], while Humphrey et al. (2019) demonstrated synergistic effects on linear growth (HAZ) when WASH and complementary feeding were combined [13]. Meiyetriani and Utomo (2025) confirmed in Indonesia that poor sanitation was independently associated with co-occurring anemia and stunting [14].

### *B. Supporting Theory*

### *B.1 K-Medoids (PAM) Clustering*

K-Medoids, introduced by Kaufman and Rousseeuw (1990), is a partition-based clustering method that selects actual data points (medoids) as cluster centres, minimising total intra-cluster dissimilarity [20]. Unlike K-Means, it is robust to outliers and extreme values, making it appropriate for aggregate provincial health data where outlying provinces (e.g., Papua highlands) cannot be removed [21][22]. The PAM algorithm iterates between assignment and medoid-swap steps until no swap decreases total dissimilarity. Manhattan (city-block) distance was selected over Euclidean distance for its lower sensitivity to multicollinearity and dimensional scale differences [20][22].

### *B.1.1 K-Medoids Method*

K-Medoids is a clustering method with a partition basis category developed by Leonard Kaufman and Peter J. Rousseeuw in 1987 [20]. The Partition Around Medoids (PAM) algorithm is the basic process of the K-Medoids method in an effort to minimize the total distance between each point and the specified medoid. This method has the same concept as K-Means, but is robust to outliers and high-dimensional data. If K-Means uses the average as the cluster center, then K-Medoids selects medoids that are the most representative actual objects in each cluster [21]. This approach makes K-Medoids more stable because it is not affected by extreme values [20]. The stages of K-Medoids are as follows [22]:

1. Initialize the medoids according to the number of clusters.
2. Calculate each object using the Manhattan distance equation, which will direct it to the cluster with the closest medoid:

$$d(x,y) = \sum_{i=1}^{n}|x_i - y_i| \tag{1}$$

   where:
   $d(x,y)$= distance between two data points
   $x$ = first data vector
   $y$ = second data vector
   $n$ = number of dimensions
   $|x_i - y_i|$ = absolute value of the difference between the two data points from component i.
3. Randomly select one object that is not a medoid.
4. Calculate the distance of each object in the cluster to the non-medoid candidate.
5. Calculate the total deviation (S). If the new total deviation is smaller than the previous total deviation, the medoid is replaced with a new candidate.
6. Repeat steps 3 through 5 until there is no change in the medoid.

### *B.1.2. Clustering Validation*

The Average Silhouette Width (ASW) is the primary internal validation criterion. For observation i, the silhouette s(i) = [b(i) − a(i)] / max[a(i), b(i)], where a(i) is mean intra-cluster distance and b(i) is mean distance to the nearest other cluster. ASW ranges from −1 to +1; values above 0.25 indicate a reasonable structure [20]. Variable contributions are assessed by Kruskal-Wallis $\eta^2$ (proportion of rank-variance explained between clusters),

$R^2$ (between-cluster variance proportion), medoid pairwise mean absolute differences (MedoidPairMeanAbs), and permutation-based ΔASW (importance of each variable to overall cluster quality).

### B.2. *Stunting*

Stunting is a malnutrition in children caused by chronic malnutrition. According to the World Health Organization (WHO, 2023), stunting is defined as a height-for-age z-score (HAZ) of less than -2 standard deviations based on the WHO Child Growth Standards [23]. This condition describes impaired linear growth associated with inadequate nutritional intake, frequent infections, and unfavorable environmental and socioeconomic factors.

Stunting not only impacts physical growth but also cognitive development and long-term productivity. Children who experience stunting are at risk of having lower intelligence than others [24]. The high prevalence of stunting in a country is an indicator of the failure of the country's health and nutrition system.

The causes of stunting are multifactorial. According to UNICEF (2022), direct causes of stunting include inadequate nutritional intake and recurrent infectious diseases, while indirect causes include poor parenting, limited food access, low access to clean water and sanitation, poor quality maternal and child services, and ANC visits [25]. According to the Indonesian Ministry of Health (2022), stunting in Indonesia is caused by low coverage of exclusive breastfeeding, poor nutrition among pregnant women, and family socioeconomic conditions [26].

### *B.3. Cross-Cluster Policy*

Cross-cluster policies for strengthening maternal nutrition emphasize the importance of synergy between the health, nutrition, and environmental sectors. Strengthening ANC services is a key strategy because it serves as a platform for specific nutrition interventions such as IFA or MMS supplementation, as well as integrated nutrition counseling. Studies show that increasing ANC visits and nutrition education can improve adherence to IFA consumption and reduce the risk of maternal anemia [6] [7]. Therefore, fully utilizing the ANC platform as part of cross-cluster policies is considered crucial for strengthening maternal nutrition interventions and reducing the prevalence of anemia.1

The effectiveness of nutrition interventions will be optimal when integrated with WASH interventions. Good sanitation and hygiene can reduce the risk of infection, diarrhea, and worm infestation, which inhibit nutrient absorption [12] [13]. Research shows that the combination of nutrition and WASH interventions has a synergistic impact on growth and hematological status, while single interventions are less effective [27] [14]. Therefore, a Nutrition-Sensitive WASH approach that integrates nutrition improvement, maternal health services, sanitation improvements, and community education needs to be a policy priority to reduce anemia prevalence and improve nutritional status sustainably [4] [5]. Thus, an effective cross-cluster strategy should combine (1) strengthening ANC and maternal nutrition services (supplementation, counseling, screening), (2) WASH interventions that reduce the burden of infection, and (3) advocacy and community empowerment to achieve higher compliance and access to nutritious food through an integrated approach, which can have greater potential to reduce the prevalence of anemia than single interventions.

## III. RESEARCH METHODOLOGY

### *A. Data Types and Sources*

Secondary data from the 2023 Indonesian Health Survey (SKI), published by the Ministry of Health of Indonesia, were used. Data are aggregated at the provincial level across all 38 provinces. Ten stunting risk indicators were selected based on theoretical relevance [4][5][9] and availability in the SKI 2023 report (Table 1). Terminology is standardised throughout: ‘maternal anemia prevalence’ refers to anemia among women aged 15–49 years; ‘maternal CED’ refers to mid-upper arm circumference below 23.5 cm; these are distinct measurement subgroups and are not used interchangeably.

### *B. Research Variable*

The risk indicators used include 10 main variables representing nutritional, maternal health, and environmental factors, namely:

1. Exclusive breastfeeding coverage (%),
2. Access to proper sanitation (%),
3. Proportion of households with unsafe drinking water (%),
4. Household wealth status (1 - 5),
5. Proportion of infants with low birth weight (LBW) (%),
6. Consumption of iron tablets by pregnant women (%),
7. Proportion of mothers with chronic energy deficiency (CED) (%),
8. Prevalence of maternal anemia (%),
9. Proportion of women aged 18 years and over who experience wasting (%), and
10. Coverage of ANC visits ⩾10 times (%).

### *C. Research Methodology*

All analyses were conducted in R (version 4.3.2) using the cluster package (v2.1.6) for PAM and factoextra (v1.0.7) for visualisation. Data were compiled in Microsoft Excel prior to import. A random seed of 42 (set.seed(42)) was used throughout for reproducibility. The analytical pipeline was: (1) descriptive analysis and outlier detection (3×IQR extreme rule) without removing anomalous data; (2) robust Median-MAD scaling: $z = (x - \text{median})/\text{MAD}$ to control extreme values; (3) VIF multicollinearity assessment; (4) ASW optimization across k=2 to 10 using Manhattan distance matrices; (5) PAM clustering fixed at k=4 utilizing the standard Partitioning Around Medoids algorithm; (6) advanced variable contribution analysis using post-hoc Kruskal-Wallis $\eta^2$, $R^2$, pairwise absolute differences between medoid values, and permutation-based ΔASW impact tracking; and (7) cluster profile interpretation and policy recommendation.

Research limitation: (1) Because all 38 provinces are included (the full population of administrative units), sampling bias does not arise. However, $n = 38$ relative to $p = 10$ warrants caution regarding cluster stability. Bootstrap-based stability testing (Jaccard index) is recommended for future work but was not performed here; (2) aggregated provincial data preclude individual-level inference (ecological fallacy); (3) mechanistically important variables i.e. malaria endemicity, helminth prevalence, dietary diversity scores — are absent from the SKI 2023 dataset, limiting causal interpretation.

## IV. RESULT AND DISCUSSION

### A. Descriptive Analysis

Prior to grouping, the data underwent initial exploration, including outlier detection using the 3 × Interquartile Range (IQR) extreme rule and multicollinearity testing using the VIF. Outlier detection identified extreme values in sanitation access and unsafe drinking water for Papua highlands provinces, confirming the necessity of robust Median-MAD scaling. VIF values ranged from 1.49 to 3.34 all below the conservative threshold of 4.0, indicating acceptable multicollinearity for clustering purposes (Table I).

TABLE I. VIF VALUE

| 1 | Variable | VIF Value |
|---|---|---|
| 2 | Exclusive breastfeeding coverage | 1.843 |
| 3 | Access to proper sanitation | 3.342 |
| 4 | Proportion of households with unsafe drinking water | 2.453 |
| 5 | Household wealth status | 2.122 |
| 6 | Proportion of infants with low birth weight (LBW) | 2.103 |
| 7 | Consumption of iron tablets by pregnant women | 2.015 |
| 8 | Proportion of mothers with chronic energy deficiency (CED) | 2.015 |
| 9 | Prevalence of maternal anemia | 1.493 |
| 10 | Proportion of women aged 18 years and over who experience wasting | 2.177 |
| 1 | Coverage of ANC visits ⩾10 times | 2.684 |

### B. Optimal Cluster

ASW was computed for k = 2 to 6. The highest ASW of 0.425 was obtained at k = 4, yielding four interpretable and operationally distinct provincial clusters. Lower ASW values at other k (k=2: 0.387; k=3: 0.361; k=5: 0.398; k=6: 0.371) confirmed k = 4 as optimal. An ASW of 0.425 reflects the expected within-country heterogeneity of a large archipelagic nation with 38 diverse provinces.

### C. K-Medoids Clustering Analysis (PAM) Results

The K-Medoids approach was conducted using robust scaling (Median-MAD) and Manhattan distance to ensure stable segmentation. The optimal number of clusters was determined using ASW, value of 0,425 obtained at k = 4. The following are the results of the K-Medoids Clustering analysis, which yielded four clusters with different stunting risk characteristics:

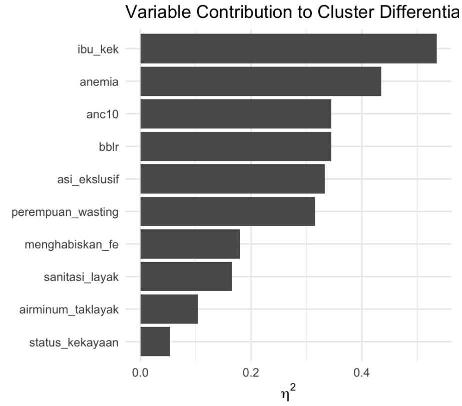

Fig. 1. Cluster Map

1. Cluster 1
   Cluster 1 includes twenty-three provinces: Banda Aceh, North Sumatra, West Sumatra, Riau, Jambi, South Sumatra, Bengkulu, Lampung, Bangka Belitung, Riau Islands, DKI Jakarta, West Java, East Java, Banten, Bali, West Kalimantan, Central Kalimantan, East Kalimantan, North Sulawesi, Southeast Sulawesi, Maluku, North Maluku, and Southwest Papua. The characteristics of this cluster indicate good WASH conditions, with an average sanitation coverage of 86,7%, ranging from 82,9% to 89,3%. Meanwhile, the proportion of households with unsafe drinking water is only 8,5%, ranging from 6,4% to 11,8%. Furthermore, exclusive breastfeeding coverage still needs to be improved, averaging 57%, and ANC coverage of only 52,1%. Furthermore, the nutritional burden of mothers and infants is relatively light, with a proportion of low birth weight (LBW) of 5,3%, maternal CED of 10,9%, maternal anemia of 4,2%, and wasting of 5,7%. Overall, this cluster depicts an area with healthy environmental conditions and a low nutritional burden, but still needs to improve the quality of ANC and exclusive breastfeeding, while maintaining good WASH quality.
2. Cluster 2
   Cluster 2 encompasses ten provinces: Central Java, Yogyakarta Special Region, West Nusa Tenggara, East Nusa Tenggara, South Kalimantan, North Kalimantan, Central Sulawesi, South Sulawesi, Gorontalo, and West Sulawesi. This cluster demonstrates good WASH conditions, with an average sanitation coverage of 86,8%, ranging from 83,42% to 89,05%. Meanwhile, the proportion of households with unsafe drinking water is only 7.2%, ranging from 6,5% to 11,5%. Furthermore, ANC coverage is also considered good, at 68,95%. However, exclusive

breastfeeding coverage still needs to be improved due to its low average of 43,7% (the lowest), low birth weight (LBW) at 7,45% (the highest), maternal CED at 26,2% (the highest), maternal anemia at 5,85%, and wasting in adult women at 8,05% (the highest). Overall, this cluster represents a region with conditions indicating a mismatch between service availability and nutritional behavior, necessitating specific behavior-based nutritional interventions and energy-protein and Fe-FA supplementation.

3. Cluster 3
Cluster 3 encompasses three provinces: West Papua, Papua, and South Papua. Characteristics of this cluster indicate very high exclusive breastfeeding coverage and FE tablet consumption, with median values of 72,8% and 71,1%, respectively. ANC coverage is also quite good at 64,7%, and sanitation is relatively good at 82,3%. However, the prevalence of anemia among women and pregnant women is very high, reaching 18,9%, maternal CED at 21,3%, wasting at 6%, and the proportion of low birth weight (LBW) at 5,3%. These conditions indicate issues with bioavailability, helminth infections, or malaria that are not measured by behavioral indicators. These include fortification, deworming, malaria control, and energy and protein adequacy in community consumption patterns.

4. Cluster 4
Cluster 4 encompasses two provinces: Central Papua and Highland Papua. This cluster's characteristics indicate very poor WASH conditions, with an average sanitation coverage of only 29,15% and a proportion of households with unsafe drinking water of only 29,45% (very high). However, exclusive breastfeeding coverage is very high, reaching 77,6% (the highest), and iron (Fe) coverage of 74,8%. Furthermore, ANC coverage is 29% (the lowest), with anemia prevalence of 4,25%, maternal CED of 11,35%, and wasting of 3,55% (the lowest). Overall, this cluster represents an area with very poor WASH environments and low ANC coverage, yet high exclusive breastfeeding rates and the lowest wasting rates. This situation demonstrates a focus on WASH and natural nutritional protection from breastfeeding behaviors, despite high-risk physical environments through the expansion of ANC to prevent LBW or complications, while maintaining high exclusive breastfeeding coverage.




.

Fig. 2. Variable Contribution

Maternal CED (Kruskal-Wallis $\eta^2$= 0.54) was the most influential factor distinguishing provincial clusters, followed by maternal anemia (Kruskal-Wallis $\eta^2$=0.44). These two indicators exhibited substantially larger effect sizes than the remaining variables, indicating that differences in maternal nutritional status and anemia prevalence were the primary drivers of cluster formation. This finding suggests that provincial disparities in stunting-related risk profiles are strongly associated with maternal health conditions, particularly nutritional deficiencies before and during pregnancy.

Table II. *Within-cluster Mean Profiles and Medoid Provinces*

| Indicator | Cluster 1 (n=23) | Cluster 2 (n=10) | Cluster 3 (n=3) | Cluster 4 (n=2) |
|---|---|---|---|---|
| Exclusive breastfeeding (%) | 57.0 | 43.7↓ | 72.8 | 77.6↑ |
| Sanitation access (%) | 86.7 | 86.8 | 82.3 | 29.2↓↓ |
| Unsafe drinking water (%) | 8.5 | 7.2 | 13.4 | 29.5↑↑ |
| Wealth status (1–5) | 3.4 | 3.2 | 2.6 | 2.1↓ |
| LBW (%) | 5.3 | 7.5↑ | 5.3 | 4.8 |
| Iron tablet consumption (%) | 55.2 | 52.8 | 71.1↑ | 74.8↑ |
| Maternal CED (%) | 10.9 | 26.2↑↑ | 21.3↑ | 11.4 |
| Maternal anemia (%) | 4.2 | 5.9 | 18.9↑↑ | 4.3 |
| Wasting ≥18 yrs (%) | 5.7 | 8.1↑ | 6.0 | 3.6↓ |
| ANC ≥10 visits (%) | 52.1 | 69.0 | 64.7 | 29.0↓↓ |
| **Medoid province** | Lampung | Central Java | West Papua | Central Papua |

### *Discussion: Cluster-Specific Policy Recommendations*

The four clusters exhibit distinct risk profiles that require differentiated policy responses. The following recommendations are explicitly aligned with existing Indonesian national programmes — including the Thousand Days of Life (HPK) programme, Gerakan Nasional Percepatan Perbaikan Gizi (Gernas PPG), Tablet Tambah Darah (TTD), Posyandu revitalisation, Jaminan Kesehatan Nasional (JKN), and Dana Otsus Papua — while identifying specific, feasible enhancements that address each cluster's identified gaps and go beyond current uniform implementation.

### *Cluster 1 — Good WASH, ANC and Breastfeeding Quality Gaps (23 provinces: Sumatra, Java, Kalimantan, North Sulawesi, Maluku, Southwest Papua)*

Cluster 1 has adequate WASH conditions and low maternal nutritional burden (anemia 4.2%; CED 10.9%), but ANC ≥10 coverage (52.1%) and exclusive breastfeeding rates (57.0%) remain below national targets. The dominant gap is service quality and behavioural uptake, not infrastructure or severe nutritional deficiency.

Policy enhancement: Within the existing HPK and Posyandu frameworks, the specific improvement is to shift from facility expansion (already relatively good) to quality and completion. Concretely: (i) integrate structured counselling checklists and appointment reminder systems (SMS/WhatsApp via kader networks) into ANC to increase ≥10-visit completion — consistent with evidence that structured ANC follow-up improves IFA adherence [6][7]; (ii) deploy certified lactation counsellors at Puskesmas level and link to existing Posyandu for breastfeeding support group formation; and (iii) use SKI data dashboards at district health offices to monitor ANC completion rates and exclusive breastfeeding coverage as accountability metrics, building on the Gernas PPG monitoring framework. These enhancements are feasible within current Puskesmas operational budgets and do not require new facility construction.

### *Cluster 2 — Good WASH, High Maternal Nutritional Burden (10 provinces: Central Java, DIY, NTB, NTT, South/North/Central Kalimantan, Central/South Sulawesi, Gorontalo, West Sulawesi)*

Cluster 2 has the highest maternal CED (26.2%), highest wasting in adult women (8.1%), and highest LBW rates (7.5%), alongside the lowest exclusive breastfeeding coverage (43.7%). WASH conditions and ANC coverage are adequate, confirming that service access is not the binding constraint — the bottleneck is nutritional status, particularly energy-protein adequacy.

Policy enhancement: The national Tablet Tambah Darah (TTD) programme provides IFA supplementation but does not address energy-protein deficits. The specific, feasible enhancement for Cluster 2 is: (i) within existing Gernas PPG and PKH (Program Keluarga Harapan) conditional cash transfers, introduce targeted food-based interventions for women of reproductive age — specifically local food fortification and Pemberian Makanan Tambahan (PMT) for underweight women, not only for children; (ii) expand CED screening from ANC first visit to community-level screening by kader at Posyandu using MUAC tape, enabling early identification of at-risk women before pregnancy; and (iii) integrate nutrition counselling on dietary diversity and energy adequacy — not only IFA compliance — into existing ANC sessions. These build on existing platforms without requiring new infrastructure and address the energy-protein gap not covered by current IFA-focused programmes [6][7].

### *Cluster 3 — High Anemia Despite Good Service Coverage (3 provinces: West Papua, Papua, South Papua)*

Cluster 3 has the highest maternal anemia (18.9%), despite high iron tablet consumption (71.1%), good ANC coverage (64.7%), and high exclusive breastfeeding (72.8%). This paradox — high IFA coverage with persistently high anemia — is inconsistent with behavioural explanations alone and strongly implicates bioavailability barriers, helminth infection, or malaria as underlying drivers not measured by the current indicator set. This is the most analytically distinctive cluster.

Policy enhancement: The existing TTD programme addresses supplementation delivery but not bioavailability or infection-related anemia loss. The specific enhancements are: (i) integrate malaria rapid diagnostic testing and treatment into routine ANC as a protocol for eastern Papua provinces, building on the existing national malaria programme (Program Nasional Penanggulangan Malaria) — this is feasible through Puskesmas and mobile health teams already operational in Papua; (ii) add preventive deworming (soil-transmitted helminth mass drug administration) for women of reproductive age into the existing community health programme cycle, consistent with WHO recommendations [5]; and (iii) introduce food-based iron bioavailability improvement (vitamin C-rich food pairing, reduction of iron inhibitors) through existing posyandu nutrition counselling. These are targeted, low-cost interventions feasible within Papua's existing health infrastructure that directly address the mechanistic gap identified by this cluster's paradoxical profile.

### *Cluster 4 — Severe WASH Deficit, Low ANC, High Breastfeeding (2 provinces: Central Papua, Highland Papua)*

Cluster 4 has the most extreme profile of all clusters: the lowest sanitation access (29.2%), highest proportion of households with unsafe drinking water (29.5%), and the lowest ANC coverage (29.0%). Paradoxically, exclusive breastfeeding (77.6%) and iron tablet consumption (74.8%) are the highest of all clusters, and wasting (3.6%)

and anemia (4.3%) are the lowest — suggesting that high breastfeeding practices provide meaningful nutritional protection within an extremely resource-constrained environment.

Policy enhancement: This cluster requires the most structurally distinct intervention package. Within the existing Dana Otsus (Special Autonomy Fund) Papua and JKN frameworks: (i) prioritise basic WASH infrastructure investment, specifically safe water access points and household-level sanitation at the village level, as the single highest-impact structural intervention, given the extreme WASH deficit that current incremental programmes have not addressed; (ii) deploy mobile ANC outreach teams (Tim Nusantara Sehat model) with extended geographic reach to increase ANC access in geographically isolated highland communities from 29% toward the national target of ≥90%; and (iii) actively protect existing breastfeeding cultural norms through health worker training and community communication — avoiding introduction of infant formula marketing in communities where traditional breastfeeding is demonstrably effective. This cluster is the one most requiring emergency-scale infrastructure investment rather than programme-quality improvement.

## V. Conclusions and Suggestions

This study applied K-Medoids PAM with robust Median-MAD scaling and Manhattan distance to segment all 38 Indonesian provinces into four distinct stunting risk clusters (ASW = 0.425). Multi-metric variable contribution analysis confirmed that maternal CED ($\eta^2 = 0.535$, rank 2.6) and maternal anemia ($\eta^2 = 0.435$, rank 1.8) are the primary drivers of provincial differentiation, while sanitation access is the structural separator for the Papua highlands. ANC coverage, exclusive breastfeeding, and wasting in adult women are important supporting discriminators.

The four clusters require substantively different, feasible policy enhancements within existing Indonesian programme frameworks: Cluster 1 needs ANC quality and breastfeeding behaviour improvement within HPK and Posyandu; Cluster 2 needs energy-protein supplementation and CED screening within Gernas PPG and PKH; Cluster 3 needs anemia-specific bioavailability interventions (malaria, deworming) within the existing Papua health infrastructure; and Cluster 4 requires urgent WASH infrastructure investment and mobile ANC expansion within Dana Otsus. These targeted, cluster-specific enhancements address the gaps not covered by uniform national programmes and provide a data-driven basis for more effective, contextually appropriate stunting reduction strategies.

## Acknowledgment

The author would like to express sincere to Universitas Mataram for their valuable support and guidance throughout this research. Appreciation is also extended to all contributors who provided insights and assistance in completing this study.

## References

[1] Sutarto, D. Mayasari and R. Indriyani, "Stunting, Faktor Resiko dan Pencegahannya," *Jurnal Agromedicine,* vol. 5, no. 1, pp. 540-545, 2018.

[2] Kementrian Kesehatan RI, "SSGI 2024: Prevalensi Stunting Nasional Turun Menjadi 19,8%," 26 Mei 2025. [Online]. Available: https://kemkes.go.id/id/ssgi-2024-prevalensi-stunting-nasional-turun-menjadi-198.

[3] Kementrian Kesehatan RI, "Prevalensi Stunting di Indonesia Turun ke 21,6% dari 24,4%," 25 Januari 2023. [Online]. Available: https://kemkes.go.id/id/prevalensi-stunting-di-indonesia-turun-ke-216-dari-244.

[4] UNICEF, "Improving nutrition outcomes with better water, sanitation and hygiene: Practical solutions for integrating WASH and nutrition programmes," UNICEF, New York, 2020.

[5] World Health Organization (WHO), "Guideline: Implementation of antenatal care recommendations on nutrition, diet and WASH," Geneva, WHO, 2021.

[6] S. M. Billah, K. K. Saha, M. A. Karim, A. N. S. Khan, A. H. Chowdhury , S. Roy and T. Ahmed, "Factors influencing maternal consumption of iron-folic acid supplements in Bangladesh: Evidence from national survey data," *BMC Pregnancy and Childbirth,* vol. 22, no. 1, pp. 1-13, 2022.

[7] M. T. Engidaw, A. D. Gebremariam and A. M. Abebe, "Effect of nutrition education interventions on adherence to iron and folic acid supplementation among pregnant women: A systematic review and meta-analysis," *PLOS ONE,* vol. 19, no. 2, 2024.

[8] BAPPENAS, "BERANDA-SDGs Indonesia," 2023. [Online]. Available: https://sdgs.bappenas.go.id/.

[9] R. Fadhilah, D. E. Putri and H. S. Amanda, "Faktor-Faktor Yang Mempengaruhi Stunting Di Indonesia: Analisis Principal Component Regression," *Jurnal Gaussian,* vol. 14, no. 1, pp. 128-138, 2025.

[10] A. R. Cahyani, E. R. Puspita, L. A. Pratiwi and M. Muhajir, "Pengelompokkan Stunting Menggunakan Metode K-Medoids Di Daerah Istimewa Yogyakarta (DIY)," *Mathematic & Applications Journal,* pp. 28-37, 2024.

[11] H. Pohan, M. Zarlis, E. Irawan, H. Okprana and Y. P. Purba, "Penerapan Algoritma K-Medoids dalam Pengelompokan Balita Stunting di Indonesia," *JUKI: Jurnal Komputer dan Informatika,* vol. 3, no. 2, pp. 97-104, 2021.

[12] T. Bekele, P. Rawstorne, B. Rahman and C. A. Gericke, "Effect of water, sanitation and hygiene interventions on nutritional status of children: A systematic review and meta-analysis of randomized controlled trials," *BMC Public Health,* vol. 20, pp. 1-15, 2020.

[13] J. H. Humprey, M. N. N. Mbuya, R. Ntozini, L. H. Moulton, R. J. Stolzfus, N. Tavengwa and The SHINE Trial Team, "Independent and combined effects of improved water, sanitation, and hygiene, and improved complementary feeding, on child stunting and anemia in rural Zimbabwe: A cluster-randomised trial," *The Lancet Global Health,* vol. 7, no. 1, pp. e132-e147, 2019.

[14] E. Meiyetriani and B. Utomo, "Association of poor sanitation with concurrent anemia and stunting among children in Indonesia: A retrospective cohort study," *Public Health Nutrition,* vol. 28, no. 3, pp. 455-463, 2025.

[15] R. S. Nurhalizah, R. Ardianto and Purwono, "Analisis Supervised dan Unsupervised Learning pada Machine Learning: Systematic Literature Review," *Jurnal Ilmu Komputer dan Informatika (JIKI),* vol. 4, no. 1, pp. 61-72, 2024.

[16] E. Alpaydin, Intoduction to Machine Learning, Fourth Edition, Cambridge: MIT Press, 2020.

[17] M. Gheisari, H. Hamidpour, Y. Liu, P. Saedi , A. Raza, A. Jalili, H. Rokhsati and R. Amin, "Data Mining Techniques for Web Mining: A Survey," *Artificial Intelligence and Applications,* vol. 1, no. 1, pp. 3-10, 2023.

[18] R. D. Vaghela and D. S. S. Iyer, "A Comparative Analysis of Clustering Algorithm," *The Electrochemical Society,* vol. 107, no. 1, pp. 2435-2443, 2022.

[19] A. E. Ezugwu, A. M. Ikotun, O. O. Oyelade, L. Abualigah , J. O. Agushaka, C. I. Eke and A. A. Akinyelu , "A Comprehensive Survey of Clustering Algorithms: State-of-The-Art Machine Learning Applications, Taxonomy, Challenges, and Future Research Prospect," *Engineering Applications of Artificial Intelligence,* vol. 110, 2022.

[20] L. Kaufman and P. J. Rousseeuw, Finding Groups in Data: An Introduction to Cluster Analysis, New York : Wiley, 1990.

[21] S. R. W. F. Erika and Suhada, "Mengkluster jumlah kabupaten/kota yang melaksanakan kawasan tanpa rokok (ktr) di 50% sekolah menurut provinsi dengan k-medoids," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer),* vol. 3, pp. 637-644, 2019.

[22] E. Rahmah, E. Haerani, A. Nazir and S. Ramadhani, "Penerapan Algoritma K-Medoids Clustering Untuk Menentukan Strategi Promosi Pada Data Mahasiswa (Studi Kasus : Stikes Perintis Padang)," *Jurnal Nasional Komputasi dan Teknologi Informasi,* vol. 5, no. 3, pp. 556-564, 2022.

[23] World Health Organization (WHO), Malnutrition: Key Facts, Geneva: WHO, 2023.

[24] C. G. Victoria, P. Christian , L. P. Vidaletti, G. Gatica-Dominguez, P. Menon and R. E. Black , "Revisiting maternal and vhild undernutrition in low-income and middle-income countries: Variable progress towards an unfinished agenda," *The Lancet,* vol. 397, no. 10282, pp. 1388-1399, 2021.

[25] UNICEF, Improving WASH and Nutrition: A System Review, New York: United Nations Children's Fund, 2021.

[26] Kementrian Kesehatan RI, Laporan Status Gizi Indonesia (SSGI) 2022, Jakarta: Direktorat Gizi Masyarakat, Kemenkes RI, 2022.

[27] L. Cameron, C. Chase and D. C. Suarez, "Impact evaluation of sanitation and hygiene interventions on maternal and neonatal health outcomes in Indonesia," *World Development,* vol. 147, p. 105637, 2021.

[28] B. Wira, A. E. Budianto and A. S. Wiguna, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 di Universitas Kanjuruhan Malang," *Jurnal Terapan, Sains, & Teknologi (RAINSTEK),* vol. 1, no. 3, pp. 54-69, 2019.