Jurnal Matematika dan Statistika serta Aplikasinya, Vol. 14, No. 1, 13-23, 2026
DOI: https://doi.org/10.24252/msa.v13i2.62887

Clustering of Regencies/Cities in NTT Province Based on Poverty Using the Ward and Fuzzy C-Means Methods

Ervi Dayana Putri[1]*, Tessy Octavia Mukhti[2]
[1,2] *Statistics Study Program, Padang State University, Padang, Indonesia*
***Corresponding author**: tessyoctaviam@fmipa.unp.ac.id

***Submission date**: 25 September 2025, **Revision**: 01 Octobert 2025, **Accepted**: 24 January 2026

ABSTRACT
Poverty remains a major problem in Indonesia's economic development. In Indonesia, East Nusa Tenggara Province has the fifth highest percentage of poor people in Indonesia as of September 2024. However, East Nusa Tenggara Province has also shown a decline in its poverty rate. According to a report from the Statistics Indonesia of Nusa Tenggara Timur Province, 18.60 percent of the province's population will be impoverished in 2024. This number is still significantly higher than Indonesia's 8.57 percent poverty rate. Additionally, East Nusa Tenggara Province's districts and cities have not benefited equally from the reduction in poverty because of notable regional differences. This study uses poverty data for East Nusa Tenggara Province in 2024 obtained from the Statistics Indonesia of Nusa Tenggara Timur Province and applies cluster analysis by comparing the Ward hierarchical cluster method and the non-hierarchical Fuzzy C-Means method to group regions based on poverty characteristics. The results of the study concluded that the Ward method with 3 clusters provided better results than the Fuzzy C-Means method. It is anticipated that the government will use these results as a foundation to concentrate more on creating more focused development plans for regions with the highest rates of poverty.

KEYWORDS
Cluster Analysis, East Nusa Tenggara, Fuzzy C-Means, Poverty, Ward Method.

1. INTRODUCTION

Poverty remains a major problem in Indonesia's economic development. The government has created various policies and programs aimed at reducing poverty rates. However, this problem remains a complex challenge in some parts of Indonesia, with groups of people still trapped in poverty.

The proportion of Indonesians living in poverty has decreased in recent years. As of September 2022, 9.57 percent of Indonesians lived in poverty; by March 2023, that number had dropped to 9.36 percent; by September 2024, it had dropped to 8.57 percent [1]. The decline in the percentage of poor people in Indonesia shows good performance. However, poverty distribution in Indonesia has not yet been realized. In terms of the percentage of poor people, Bali has a lower percentage than other provinces, at 3.80 percent as of September 2024, while Papua Pegunungan Province has the highest poverty rate at 29.66 percent. This shows that there is a high level of inequality between provinces in Indonesia.

East Nusa Tenggara Province has the fifth highest percentage of poor people in Indonesia as of September 2024[2]. Based on a publication by the Statistics Indonesia of Nusa Tenggara Timur Province, the percentage of poor people in East Nusa Tenggara Province as of September 2024 was 19.02 percent. This figure is still far above the poverty rate in Indonesia, which stands at 8.57 percent. In addition, poverty reduction has not been evenly distributed across districts/cities in East Nusa Tenggara Province, due to significant regional disparities. Differences in geographical conditions in various regions pose a challenge to achieving equitable poverty reduction. For example, in 2024, the percentage of poor people in Central Sumba was 30.84 percent, much higher than in Kupang City, which was only 8.24 percent[3]. In understanding these patterns of disparity, a

data analysis approach is needed that can group regions based on similar poverty characteristics. One approach that can be considered in addressing this issue is cluster analysis.

Hierarchical and non-hierarchical methods are the two primary approaches used in cluster analysis. One example of a hierarchical approach is Ward's method, which reduces the amount of information lost when objects are combined into clusters. [2]. In its conventional form, Ward's method begins with the observed objects as single units, namely clusters. These clusters are then gradually merged until a cluster containing all the objects under study is formed [3]. In addition, non-hierarchical methods such as Fuzzy C-Means (FCM) allow data to be members of several clusters at once with varying degrees of membership, generally on a scale of 0 to 1 [4]. This shows that each piece of data can have a different level of membership for each group.

Several previous studies have applied the Ward and Fuzzy C-Means methods in various analysis topics. For example, a study comparing the Ward and Fuzzy C-Means methods in analyzing Indonesia's economic growth concluded that the Ward method was better suited to the economic growth data [5] In addition, another study discussing the grouping of districts/cities based on village potential indicator data used a comparative method and proved that the Ward method was also the best method for village potential indicator data [6].

Inaccuracies in the selection of clustering methods can result in erroneous policy recommendations, especially in areas with high heterogeneity. By comparing hierarchical (ward) and non-hierarchical (fuzzy c-means) clustering, this study aims to identify areas that require special attention in order to develop policy strategies to reduce poverty, particularly in areas with the highest poverty rates in East Nusa Tenggara.

2. LITERATURE REVIEW

2.1 Data Standardization

When research data contains several variables with different units, data standardization is usually used to address this issue [6]. Z-score standardization is one of the most commonly used data standardization methods. Equation (1) is the standardization formula using z-score:

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

where Z_i is the i data point's z-score value, x_i is the i observation, $\bar{x}$ is the sample mean, and s is the standard deviation.

2.2 Multicollinearity and Kaiser-Meyer-Olkin (KMO) Tests

An assumption test used to assess an analysis's viability is the Kaiser-Meyer-Olkin (KMO) test. The KMO test value ranges from 0 to 1. If the KMO value is < 0.5, it indicates that the sample is not representative of the population, and vice versa[9]. Multicollinearity is a condition where there is correlation between independent variables. The presence of multicollinearity in the data is a violation in cluster analysis. Multicollinearity can be identified by VIF. Multicollinearity issues are indicated if the VIF value is greater than 10, and vice versa [7]. The VIF value can be determined using Equation (2):

$$VIF_j = \frac{1}{1 - R_j^2} \tag{2}$$

where R_j^2 is the coefficient of determination between x_j as a response to all other independent variables in the model.

The Kaiser-Meyer-Olkin (KMO) test is an assumption test that assesses an analysis's viability. The value of the KMO test is between 0 and 1. A KMO value of less than 0.5 suggests that the sample is not representative of the population, and vice versa [8].

2.3 Ward Method

A popular hierarchical technique in cluster analysis is the Ward method, which is based on dispersion. Its main objective is to minimize variance within clusters. This approach differs from other hierarchical methods in that it computes the distance between clusters using the principle of variance analysis. The advantage of the Ward method lies in its greater efficiency and tendency to form small clusters [7]. The steps for the Ward method are as follows:

1. Using the Silhouette Coefficient, define k (the number of clusters to be produced).

2. Calculate the distance between clusters (for example, using Euclidean Distance), where each object is considered as one cluster so that n = N. The calculation of Euclidean distance is shown in equation (3):

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{3}$$

where, D_{ij} = Euclidean distance; $x_i = x$ coordinate of object i; $y_i = y$ coordinate of object i; $x_j = x$ coordinate of object j; $y_j = y$ coordinate of object j;

3. Determine 2 of N clusters through the smallest SSE (Sum of Square Error) value as the first cluster. SSE is obtained through equation (4):

$$SSE = \sum_{i=1}^{n} (x_i - \bar{x})' (x_j - \bar{x}) \tag{4}$$

where, x_i = column vector of object i, x_j = column vector containing the average of objects in the cluster; and n = number of objects.

4. To reduce heterogeneity, examine the $N-1$ set. As a result, N clusters are (systematically) reduced by $N-1$.

5. Until every object has been sorted into one or more clusters, repeat steps 2 and 3.

2.4 Fuzzy C-Means Method

The Fuzzy C-Means method clusters data based on membership levels ranging from 0 to 1 [8]. The degree of similarity of data to a cluster is reflected in its membership value. The steps of this method are as follows:

1. First, determine the initial parameters, namely the number of clusters (c), fuzzy weight (m), initial objective value (p_0), minimum error (ε), initial iteration (t), and maximum number of iterations (T).

2. Create a membership matrix $\mathbf{U} = \{u_{ij}\}$ sized $c \times n$ randomly.

3. Calculate the cluster center (V), where the elements of matrix $\mathbf{V}$ are $\{v_i\}_c$ obtained using equation (5):

$$v_i = \frac{\sum_{j=1}^{n} (u_{ij})^m x_j}{\sum_{j=1}^{n} (u_{ij})^m} \tag{5}$$

4. Calculate the change in membership matrix elements using equation (6):

$$u_i = \frac{1}{\sum_{jk=1}^{c} \left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{m-1}}} \tag{6}$$

Next, calculate the objective value:

$$J(U,V) = \sum_{i=1}^{c} \sum_{j=1}^{n} (u_{ij})^m {u_{ij}}^2 \tag{7}$$

5. Stop if the difference in objective values $|J_t - J_{t+1}| <$ smallest error (ε) or ($t > T$) otherwise $t = t+1$, return to step 2.

2.5 Optimal Cluster Selection

2.5.1 Sillhouette Coefficient Method

The Sillhouette Coefficient is a metric used to assess how similar the data in a group is. The closer the value is to 1, the better the grouping quality [9]. The calculation formula for this method is shown in equation (8):

$$SC = \sum_{i-1}^{n} s(i) \tag{8}$$

with $s(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}}$

2.5.2 Elbow Method

The Elbow method helps determine the optimum cluster by evaluating the percentage comparison of clusters in the form of elbow points [9]. The selection of the optimum cluster can be done by calculating the Sum of Square Error (SSE) using equation (9):

$$SSE_k = \sum_{l=1}^{p} \sum_{i=1}^{m_k} |x_i - \bar{x}_{kl}|^2 \tag{9}$$
$$SSE_{total} = \sum_{k=1}^{K} SSE_k$$

where x_i is the i object and $\bar{x}_{kl}$ is the center of the k cluster of the l variable.

2.5.3 Xie-Beni Index

The Xie-Beni Index can be used to determine the ideal number of clusters in Fuzzy C Means. The purpose of the Xie-Beni Index is to calculate the ratio of total cluster variation (S) and the separation between clusters (d_{min}) [10]. The Xie Beni Index is expressed by equation (10):

$$XB_{(c)} = \frac{S}{d_{min}} \tag{10}$$
$$S = \frac{\sum_{i=1}^{n}\sum_{k=1}^{c}(u_{ik})^m d_{ik}^2(x_i, v_k)}{n}$$
$$d_{min} = min_{i=k}\, d_{ik}^2(x_i, v_k)$$

where, $d_{min}(x_i, v_k)$ is the distance of the data to the cluster center, $d_{ik}(v_i, v_k)$ is the distance from cluster center v_k to cluster center v_i.

2.6 Best Clustering Evaluation

2.6.1 Calinski-Harabasz Index

The ratio between the Sum of Square Within Clusters (SSW), which represents compactness, and the Sum of Square Between Clusters (SSB), which denotes separation, is the Calinski-Harabasz (CH) validity index. Next, this ratio is multiplied by a normalization factor, which is calculated by dividing the number of clusters minus one by the difference between the total number of data and the number of clusters [11]. The cluster outcomes are better when the CH value is higher. Equation's (11) formula can be used to determine the CH index value:

$$CH = \frac{trace\ (SSB)}{trace\ (SSW)} \times \frac{N-k}{k-1} \tag{11}$$

where k is the number of clusters and N is the total number of objects examined.

2.6.2 Sillhouette Coefficient

The cohesion and separation techniques are combined in this method. The separation value is derived from the average distance between each object in the cluster and the closest other cluster, whereas the cohesive value is computed by taking into account every object within a single cluster [12]. Cluster evaluation using this method can be calculated using equation (12):

$$sill\ (c) = sil\,(k)\,\frac{1}{|k|}\sum_{i=1}^{k} sil\,(c_i) \tag{12}$$

where, $sil(k)$ = silhouette value for all clusters; k = number of clusters k; $sil(c)$ = average silhouette value.

2.6.3 Dunn-Index

The Dunn Index is one of the validation methods based on internal criteria. This method calculates the minimum value of the comparison of the distance between two clusters as separation with compactness, which is the maximum value of the diameter [13]. This method can be calculated using equation (13):

$$I = \frac{\text{min.separation}}{\text{max.diameter}} \tag{13}$$

where, the higher the DI value, the better the cluster quality.

3. METHODOLOGY

All districts and cities in East Nusa Tenggara Province in 2024 are covered by the poverty data used in this study.The secondary data were sourced from the official website of Statistics Indonesia of Nusa Tenggara Timur Province, available at https://ntt.bps.go.id. An outline of the research data variables that will be utilized is given in **Table 1**: The method used in this

Table 1. Variables of the Study

Variable	Scale	Description
X_1	Numeric	Percentage of Poor Population
X_2	Numeric	Percentage of Households with Access to Clean Water
X_3	Numeric	Human Development Index
X_4	Numeric	Percentage of Monthly Per Capita Expenditure on Food
X_5	Numeric	Literacy Rate

study compares the Ward method with fuzzy c-means to group regencies and cities in East Nusa Tenggara Province using R software. The steps of the analysis are as follows:

1. Standardization of research data
2. Multicollinearity and KMO tests
3. Clustering using the Ward method
4. Clustering using the fuzzy c-means method
5. Evaluation of cluster results (Ward and fuzzy c-means)
6. Identification of cluster characteristics
7. Cluster mapping

4. RESULT & DISCUSSION

4.1 Data Standardization

Initial exploration of the data revealed that the variables in this study had different units of measurement (heterogeneous). To ensure that the measurement scales were equivalent, a standardization process was carried out using the z-score method. The results of the data standardization process can be seen in **Table 2**:

Table 2. Poverty Data After Standardization

Region	X_1	X_2	X_3	X_4	X_5
West Sumba	0.96	-0.36	-0.20	-0.04	-1.75
East Sumba	1.04	-1.45	0.42	0.44	0.21
⋮	⋮	⋮	⋮	⋮	⋮
Malaka	-0.94	0.34	-0.82	-0.08	-0.08
Kupang City	-1.79	0.62	3.55	-3.33	1.13

4.2 Assumption Test

To ensure the validity of the research data, the Kaiser-Meyer-Olkin (KMO) sample adequacy test must be conducted. Details of the results of both tests are presented in **Table 3**: The chosen sample is representative (fit for usage) according to an assessment

Table 3. Results of Multicollinearity and KMO Assumption Tests

Variable	**KMO**	**VIF**
X_1	0.79	2.57
X_2	0.76	2.70
X_3	0.77	2.36
X_4	0.74	2.49
X_5	0.91	2.65

of the KMO assumption ≥ 0.5. It is evident from the computations in **Table 3** shows that the KMO value for each variable is above 0.5 and the VIF value is below 10.

4.3 Ward Method

The initial stage of the Ward method begins with determining the optimum number of clusters using the Elbow and Silhouette methods. The optimum number of clusters formed is observed in **Table 1**:

Based on **Figure 1a** (Elbow method), it can be seen that the decrease in the Within Sum of Squares (WWS) value begins to slow down (Elbow point) when there are 3 clusters. After 3 clusters, the decrease in WWS is no longer significant. Meanwhile, **Figure 1b** (Silhouette method) shows that the third cluster has the best cluster quality, as it produces the highest Silhouette value. Considering both analyses, the optimum number of clusters to be considered is 3 clusters.

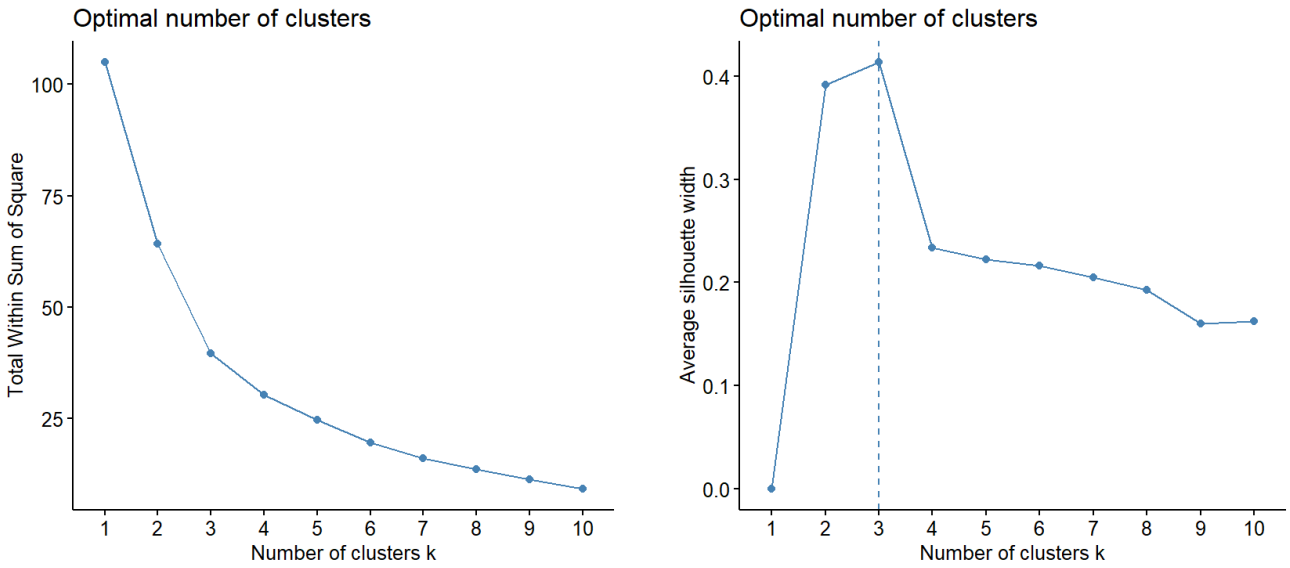
4.4 Fuzzy C-Means Method

The Xie-Beni Index is used by the Fuzzy C-Means approach to calculate the ideal number of clusters generated. **Table 4** displays the findings of figuring out the ideal number of clusters:

Table 4. Determination of Optimum Clusters Using the FCM Method

Method	**Number of Clusters**	
	2	**3**
Xie-Beni Index (XBI)	0.30	0.86

According to the results in the **Table 4**, cluster 2 has the lowest Xie-Beni Index value. For the Fuzzy C-Means approach, two clusters is the ideal number.



(a) Elbow Method **(b)** Sillhouette Method

Figure 1. Plot of the Number of Optimum Clusters Using the Ward Method

4.5 Evaluation of Results

After determining the number and distribution of clusters for each district/city, a validation test was conducted on each method to determine the best cluster. The evaluation methods used were Silhouette, Calinski Harabasz Index, and Dunn Index. The test results are presented in **Table 5**:

Table 5. Best Cluster Validation Results

Method	Number of Cluster	SC	CH	DI
Fuzzy C-Means	2	0.35	13.49	0.21
Ward	3	0.41	15.77	0.39

From the results of the cluster validation test in **Table 5**, it was found that the Ward method with 3 clusters had higher Silhouette coefficient and Dunn Index values than the Fuzzy C-Means method. This shows that the Ward method with 3 clusters is the best method for this poverty case.

4.6 Interpretation and Characteristics of Clusters

The centroid (average variable value) reflects the unique characteristics of each cluster. This centroid is used to identify patterns, analyze relationships between members (districts/cities) within clusters, and compare specific profiles between clusters. The centroid results are presented in **Table 6**:

Table 6. Centroid of Each Cluster

Cluster	X_1	X_2	X_3	X_4	X_5
1	27.5	75.6	64.9	58.8	89.7
2	18.6	91.0	68.7	56.4	96.3
3	8.24	93.7	83.2	44.6	98.7

The results in **Table 6** show that the clustering successfully divided the districts/cities in East Nusa Tenggara Province into three groups based on their respective characteristics, namely:

1. **Cluster 1**: Cluster 1 is characterized by a relatively high poverty rate, reflected in the percentage of poor people reaching around 27.5 percent. The quality of basic services is still limited, with access to clean water at 75.6 percent and an HDI value in the low-to-medium category. Areas in this cluster also show a high proportion of food expenditure, namely 58.8 percent, which indicates limited purchasing power among the population. However, the literacy rate reaches 89.7 percent, indicating that the population's basic education capacity is quite good.

2. **Cluster 2**: Cluster 2 has a low-to-medium poverty rate, with approximately 18.6 percent of the population living in poverty. Access to basic facilities is good, as indicated by 91 percent access to clean water and a literacy rate of 96.3 percent. The HDI value is in the middle category and is better than Cluster 1. Food expenditure is still quite high, at 56.4 percent, but has shown a decline in line with improving socio-economic conditions. In general, this cluster describes areas that are moving towards more prosperous conditions.

3. **Cluster 3**: Cluster 3 is a group of regions with the highest welfare, characterized by a very low percentage of poor people, namely around 8.24 percent. The quality of human development is also high, reflected in an HDI value of 83.2 and an almost perfect literacy rate of 98.7 percent. Access to clean water reaches 93.7 percent, meaning that basic services are optimal. The proportion of food expenditure is only 44.6 percent, the lowest in all clusters, indicating better purchasing power and economic conditions. Thus, this cluster represents the most advanced and prosperous regions.

The division of regencies/cities included in the formed cluster groups is visualized in **Figure 2**:

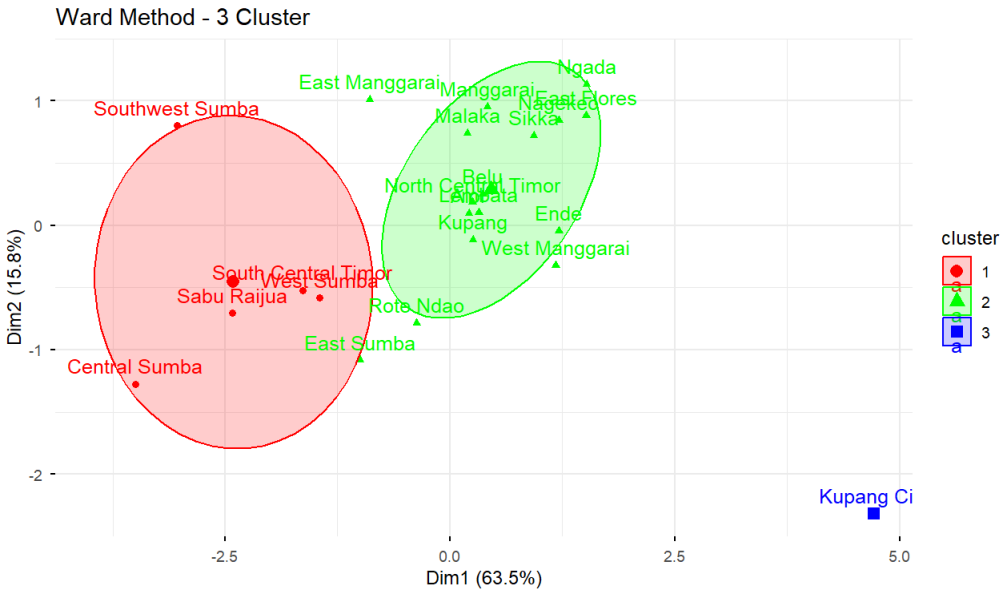


Figure 2. Visualization of Poverty Grouping by Regency/City in NTT using the Ward 3 Cluster Method

Based on **Figure 2**, it can be seen that the districts/cities included in cluster 1 consist of West Sumba, South Central Timor, Central Sumba, Southwest Sumba, and Sabu Rijua. Meanwhile, cluster 2 consists of East Sumba, Kupang, North Central Timor, Belu, Alor, Lembata, East Flores, Sikka, Ende, Ngada, Manggarai, Rote Ndao, West Manggarai, Nagekeo, East Manggarai, and Malaka. In cluster 3, only the city of Kupang is included in this cluster.

4.7 Cluster Mapping

The cluster mapping results are visualized in the following image:

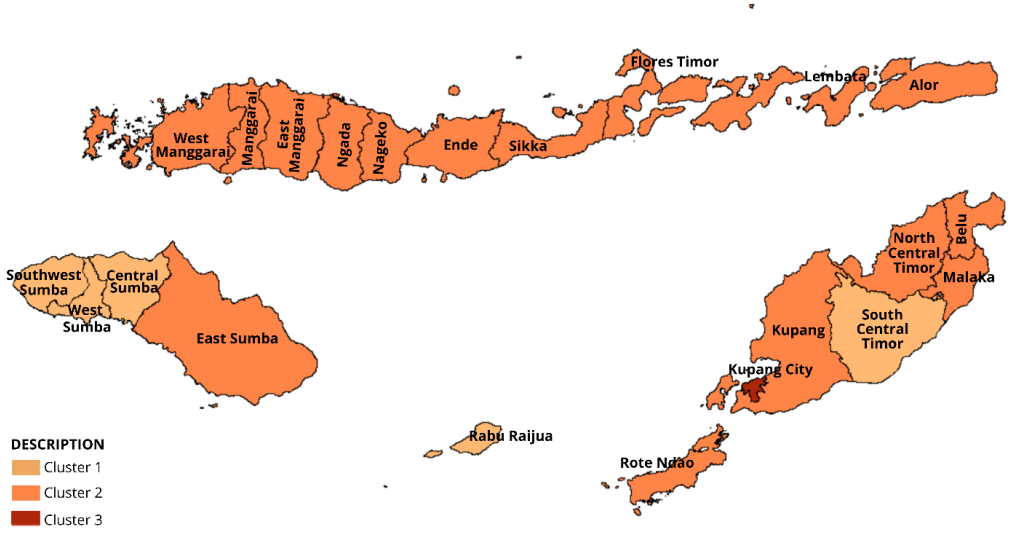


Figure 3. Mapping of Poverty by District/City in East Nusa Tenggara Province in 2024

The mapping in **Figure 3** shows a pattern of disparity between regions in East Nusa Tenggara. Almost every district/city in East Nusa Tenggara Province is in cluster 2 with a moderate poverty rate. Areas in the western part of Sumba Island tend to be in cluster 1 with the highest poverty rates, such as in Central Sumba. Meanwhile, the majority of areas on Flores Island are in cluster 2. In addition, Timor Island has regions with quite diverse poverty levels. The majority of regions on Timor Island are in cluster 2, but Kupang City is the only region in cluster 3. Thus, regions on Timor Island are beginning to transition to lower poverty levels, although there are still regions with the highest poverty levels, such as South Central Timor Regency.

4.8 Discussion

The findings in this study indicate that the Ward method is a better approach than the fuzzy c-means method. This result is proven, where research conducted by [14] shows that Ward's performance is more optimal and superior in stability compared to Fuzzy C-Means. The clustering pattern that is formed illustrates the disparity of poverty between regions in NTT. This phenomenon is in line with research that confirms that the socio-economic structure in NTT Province varies between islands and districts/cities. A study conducted by [15] concluded that Kupang and Ende were the regencies/cities with the lowest poverty rates, while Central Sumba was the region with the highest poverty rate. This study produced the same findings as this study, where Kupang was in cluster 3 with a low poverty rate, while Central Sumba was in cluster 1 with the highest poverty rate. Overall, the results of this clustering provide a spatial picture of poverty in NTT. These findings reinforce the argument of previous research by [16] which states that government program interventions cannot be homogeneous and should not be.

5. CONCLUSION

This study classified poverty conditions in districts and cities in East Nusa Tenggara Province using two clustering approaches, the Ward Clustering method and Fuzzy C-Means. The validation results using the Silhouette Coefficient, Calinski-Harabasz Index, and Dunn Index concluded that the Ward method with 3 clusters provided the best results. This method separates areas with low poverty levels in cluster 3 (Kupang City), areas with moderate poverty levels in cluster 2 (East Sumba, Kupang, North Central Timor, Belu, Alor, Lembata, East Flores, Sikka, Ende, Ngada, Manggarai, Rote Ndao, West Manggarai, Nagekeo, East Manggarai, and Malaka) to areas with high poverty levels in cluster 1 (West Sumba, South Central Timor, Central Sumba, Southwest Sumba, and Sabu Rijua). The city of Kupang, as a coastal urban center, forms its own cluster with the lowest poverty rate, while the regencies on Sumba Island mostly show higher poverty rates. These results provide a strong basis for the argument that government program interventions cannot be homogeneous and should be tailored to the dimensions of poverty based on the clusters that have been formed. Given the limitations of the study, which only used cross-sectional data for one year, the policy implications of this research need to be considered within the scope of the data and methods used.

However, the results of regional grouping in NTT can be used as a reference for the government in designing policies that can reduce poverty rates, such as providing capital assistance and entrepreneurship training, especially for areas on Sumba Island, to support poverty reduction efforts in East Nusa Tenggara Province. The strengths of this study lie in the comparative application of two clustering methods and the use of a multi-evaluation index, which provides a reliable classification framework. Further studies are recommended to apply other clustering methods, such as K-Medoids and Single Linkage, and to include additional supporting variables to generate more meaningful and in-depth policy insights.

REFERENCES

[1] BPS RI, "Profil kemiskinan di indonesia maret 2025," BPS RI, Laporan Statistik 63, 2025.

[2] Y. I. Harnanto, A. Rusgiyono, and T. Wuryandari, "Penerapan analisis klaster metode ward terhadap kabupaten/kota di jawa tengah berdasarkan pengguna alat kontrasepsi," *J. Gaussian*, vol. 6, no. 4, pp. 528–537, 2017.

[3] A. Młodak, "K-Means, Ward and Probabilistic Distance-Based Clustering Methods with Contiguity Constraint," *J Classif*, vol. 38, no. 2, pp. 313–352, Jul. 2021.

[4] A. M. A. R. Bahri, "Penentuan cluster terbaik menggunakan metode fuzzy c-means dengan analisis komponen utama," Makassar, 2024.

[5] I. K. M. M. Ningsih and A. W. Wijayanto, "Komparasi metode Clustering Provinsi pada provinsi di Indonesia Berdasarkan Pertumbuhan Ekonomi Tahun 2022," *Komputika*, vol. 13, no. 1, pp. 103–112, Apr. 2024.

[6] I. R. Akolo, A. R. Pratama, and A. Nadjamuddin, "Perbandingan Metode Fuzzy C-Means dan Ward Pada Pengelompokkan Desa Berdasarkan Indikator Potensi Desa," *Euler J. Ilm. Mat. Sains dan Teknol.*, vol. 11, no. 2, pp. 192–204, Nov. 2023.

[7] M. Y. Matdoan, A. M. Balami, and M. W. Talakua, "Implementasi metode Ward clustering untuk klasterisasi penduduk yang memiliki jaminan kesehatan pada kabupaten dan kota di Provinsi Maluku," *MIMS*, vol. 23, no. 1, p. 69, Apr. 2023.

[8] R. Siringoringo and J. Jamaluddin, "Peningkatan Performa Cluster Fuzzy C-Means pada Klastering Sentimen Menggunakan Particle Swarm Optimization," *JTIIK*, vol. 6, no. 4, pp. 349–354, Jul. 2019. [Online]. Available: https://jtiik.ub.ac.id/index.php/jtiik/article/view/1090

[9] T. Rahmawati, Y. Wilandari, and P. Kartikasari, "Analisis Perbandingan Silhouette Coefficient dan Metode Elbow Pada Pengelompokkan Provinsi di Indonesia Berdasarkan Indikator IPM dengan K-Medoids," *J.Gauss*, vol. 13, no. 1, pp. 13–24, Aug. 2024.

[10] D. R. Ningrat, D. Asih, I. Maruddani, and T. Wuryandari, "Analisis cluster dengan algoritma k-means dan fuzzy c-means clustering untuk pengelompokan data obligasi korporasi," *J. Gaussian*, vol. 5, no. 4, pp. 641–650, 2016.

[11] D. A. Saidah, R. Santoso, and T. Widiharih, "Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Kesehatan Lingkungan Menggunakan Metode Partitioning Around Medoids dengan Validasi Indeks Internal," *J.Gauss*, vol. 11, no. 2, pp. 302–312, Aug. 2022.

[12] S. Paembonan and H. Abduh, "Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat," *JIIT*, pp. 48–54, Sep. 2021.

[13] A. Nurzida, I. T. Utami, and M. Y. Rochayani, "Perbandingan Metode Fuzzy C-Means Dan Gustafson-Kessel dalam Penentuan Cluster Tingkat Risiko Penularan Tuberculosis Terhadap Penyakit di Jawa Timur," *J.Gauss*, vol. 13, no. 2, pp. 373–382, Dec. 2024.

[14] M. R. Paul and T. Handhayani, "Perbandingan fcm dan hierarchical clustering pada data ubi jalar indonesia," *J. Ilmu Komput. Sist. Inf.*, vol. 13, no. 2, 2025.

[15] M. Rahmah, "Pasar modal," *Otoritas Jasa Keuang.*, vol. 1, no. 1, pp. 1–5, 2019. [Online]. Available: https://accurate.id/ekonomi-keuangan/pengertian-pasar-modal-lengkap/

[16] L. Y. Ledhe, "Analisis Kemiskinan Multidimensi di Nusa Tenggara Timur Tahun 2023," *JSTAR*, vol. 4, no. 1, pp. 40–57, Jun. 2024.