JURNAL INOVASI DAERAH
Volume 02 Nomor 02, Desember 2023
pp.226-244, DOI:10.56655/jid.v2i2.132

CLUSTERING ANALYSIS OF MULTIDIMENSIONAL POVERTY IN CENTRAL JAVA PROVINCE, INDONESIA

Yuniasih Purwanti*
Graduate School of International Cooperation Studies (GSICS), Kobe University, Jepang
Email: yuniasihpurwanti01@gmail.com

***Penulis Korespondensi**

ABSTRAK

Kemiskinan masih menjadi masalah serius di Indonesia. Pada tahun 2020 Badan Pusat Statistik mencatat persentase penduduk miskin di tanah air sebesar 9,78%, dimana 50% dari total penduduk miskin berada di Pulau Jawa. Angka kemiskinan di Jawa Tengah yang masih di atas rata-rata nasional menjadi tantangan bersama, sehingga pengentasan kemiskinan menjadi prioritas. Kajian ini mengukur kemiskinan berdasarkan klaster berupa berbagai deprivasi yang dialami penduduk yang mencakup 35 kabupaten/kota di Provinsi Jawa Tengah. Variabel yang digunakan dalam penelitian ini terdiri dari 16 indikator kemiskinan berdasarkan Kriteria dan Pendapatan Masyarakat Miskin dan Kurang Mampu dari Kementerian Sosial. Analisis cluster menggunakan Fuzzy C Means (FCM) digunakan untuk memperoleh karakteristik kemiskinan dengan memilih cluster yang optimal. Setiap kota/kabupaten yang mempunyai indikator kemiripan dengan kota/kabupaten lain maka akan terdapat satu klaster. Klaster-klaster tersebut sangat penting dalam memahami faktor-faktor penentu kemiskinan dan program pengentasan kemiskinan. Dari hasil klasterisasi, model klaster yang terbaik adalah empat klaster dan menunjukkan bahwa pengeluaran non-makanan, kecukupan air minum, akses terhadap fasilitas sanitasi, tingkat pengangguran terbuka, dan kepemilikan televisi merupakan indikator yang paling berkorelasi terhadap kemiskinan.

Kata Kunci: *kemiskinan multidimensi, klustering, fuzzy c-means klustering*

ABSTRACT

In Indonesia, poverty continues to be a major issue. According to Statistics Indonesia, there were 9.78% of the population living in poverty in 2020, with Java Island accounting for 50% of the nation's total poor people. Furthermore, poverty reduction is the primary concern due to Central Java's high poverty rate, which is still higher than the national average, so it has become a shared challenge. This study measured poverty by clusters based on a variety of deprivations that residents of 35 regencies or municipalities in Central Java Province experienced. Additionally, 16 poverty indicators based on the criteria and income of the poor and underprivileged from the Ministry of Social Affairs comprised the variables used in this study. Moreover, by selecting the optimal cluster, characteristic poverty was obtained employing fuzzy C-means (FCM) as cluster analysis. In addition, each municipality/regency that shares a similarity indicator with another municipality/regency was categorized into one cluster. The clusters were fundamental to understanding the determinants of poverty and poverty alleviation programs. According to the clustering results, there were four clusters considered the best cluster, and it demonstrated that the indicators most associated with poverty were non-food expenditure, drinking water adequacy, access to sanitation facilities, the open unemployment rate, and television ownership.

Keywords: *multidimensional poverty, clustering, fuzzy c-means clustering*

INTRODUCTION

The United Nations (UN) members agreed to the Sustainable Development Goals (SDGs) on September 25, 2015, as a set of global development objectives. The Millennium Development Goals (MDGs) and SDG agendas state that the major objective of development initiatives is to reduce poverty [1]. Moreover, as a commitment from the Indonesian Government to eradicate extreme poverty by 2030, the National Medium-Term Development Plan (*Rencana Pembangunan Jangka Menengah Nasional* (*RPMJN) 2005–2025*) was established by the government [2]. The main target of the development plan is to synergize, in a multi-sectoral manner, both the central and regional governments, ensuring that the policy directions of the work plan are mutually supportive. Nationally, Indonesia's percentage of poor people decreased significantly from 2001 to 2020. According to Statistics Indonesia (2020a), in March 2020, 9.78% (26,42 million people) of the country's population were categorized as living below the poverty line. Additionally, the rural areas of Indonesia have been the major contributors to poverty due to the disparity in the distribution of development between western and eastern Indonesia. Furthermore, it appears that poverty is a more serious issue on islands other than Java since it is more prevalent in eastern Indonesian regions such as Papua and West Papua. In addition, the percentage of poor people is higher outside of Java. However, the absolute number of poor people is highest in Java due to its large population [4].

The region with the highest poverty rate on Java Island from 2010 to 2020 was the Special Region of Yogyakarta Province, with average poverty rates of 15.4%. Central Java came in second with an average poverty rate of 14.2%, followed by East Java, West Java, Banten, and DKI Jakarta, which had respective average poverty rates of 13.8%, 10.1%, 6.5%, and 2.4% (Statistics Indonesia, 2020a). Due to the fact that poverty levels >10% are classified as hardcore poverty by Castaneda Aguilar et al. (2016) and are higher than the national poverty average (9.78%), poverty levels above 10% can be categorized as high. Additionally, the high poverty rate in Central Java coexists with the high inequality (measured by the GINI ratio) of the population. It was indicated by the fact that 19 regencies had poverty rates above 10%, while 7 regencies and 6 municipalities had rates below 10% [4]. Furthermore, Kebumen Regency had the highest poverty rate in Central Java at 17.59% (211,09 thousand people), while Semarang Municipality had the lowest poverty rate in Central Java at 4.34% (9,690 thousand people). Therefore, it suggests that the Central Java Provincial Government has not succeeded in achieving the goal set forth in the 2020–2024 National Medium-Term Development Plan (RPJMN), which designated a 7–6.5% decrease in the number of the poor (18.34 million to 19.75 million poor people) [2], [6].

A number of factors, including level of education, social conditions, and employment opportunities, contribute to poverty. Every year, Indonesia sees a constant decrease in the number of the poor. The government had implemented several anti-poverty programs to eradicate poverty. One of the government programs is the National Team for the Acceleration of Poverty Reduction (TNP2K), which provides initiatives such as unconditional cash transfers, subsidized rice, school operation assistance programs, and targeted scholarships [7]. Moreover, a monetary approach, such as income or expenditure, is still used to measure poverty as of the present day, including in Indonesia. Thus, poverty is considered a single phenomenon (unidimensional), whereas poverty is not only determined by a monetary measurement but can also be determined by the food consumption level required to meet caloric intakes, prevalent consumption patterns, or the price of a food basket as a basic need [8], [9]. Dissatisfaction with the measurement of monetary poverty began to emerge when actual data revealed that a country's economic development rate did not immediately lower the prevalence of poverty and unemployment in developing nations. Therefore, alternative measurements were needed to complement conventional measurements (monetary approaches) [10].

Globally, the World Bank continues to employ income or consumption as indicators of poverty measurement, defining extreme poverty as living expenses below US$1.9 per day and moderate poverty as living expenses below $3.10 per day [11], [12]. Moreover, basic needs can be met by spending cash from a source of income [13]. This approach allows for the separation of the population into poor and non-poor groups based on poverty lines, which can be either absolute or relative poverty [14]. Additionally, poverty was defined as an individual's inability to access well-being as well as their inability to purchase basic goods or services [15], [16]. To address this issue, recent researchers have consistently combined monetary and non-monetary measurements, such as Battiston et al. (2013), Baulch and Masset (2003), Duclos et al. (2006), McKay & Lawson (2003), Scott (2002), Whelan (2004), and Yu (2013). Furthermore, Sen (1976) suggested two steps in the development of poverty measurement: the identification and aggregation of the poor. Since then, further measures of poverty have been introduced, one of which was the Foster-Greer-Thorbecke (FGT) [24]. A new type of multidimensional poverty measure known as the "Multidimensional Poverty Index" (MPI) was developed by Alkire and Foster (AF) and combined with the FGT measure [24], [25]. In addition, the MPI comprises three dimensions, including education, health, and standard of living, each of which has equal relative weights, as well as ten indicators within a dimension, each of which has equal weights.

Previous research on multidimensional poverty used various datasets in several countries in addition to the MPI released by the United Nations Development Programme (UNDP) to analyze it. Those previous studies are by Bérenger (2017) on Egypt and Jordan, the United Nations (2018) on Comoros, Alkire et al. (2021) on multidimensional poverty in India, and Yu (2013) on China. Furthermore, utilizing the China Health and Nutrition Survey (CHNS) 2000–2009 and the AF methodology, Yu's (2013) research examined multidimensional poverty. The research discovered that this method was able to evaluate both income reduction and other dimensions. The research also discovered that China has a relatively low rate of multidimensional poverty, with only 5% of households suffering poverty in two or more selected dimensions. Despite the fact that multidimensional poverty studies have been developed and are beginning to be employed as official poverty estimates in a number of nations, there are currently few studies on the subject in Indonesia. Additionally, researchers who have investigated it include Alkire et al. (2019); Ballon and Apablaza (2012); Hanandita and Tampubolon (2016); Indriani (2019); OPHI (2011, 2019); and Thomas et al. (2012). Moreover, longitudinal data from the Indonesia Family Life Survey (IFLS) were employed by Thomas et al. (2012) to estimate the poverty index. However, the IFLS data did not accurately represent the Indonesian population due to the fact that the sample was mostly from western Indonesia, which was relatively more developed than eastern Indonesia [35].

Meanwhile, Alkire and Santos employed IDHS to examine the MPI in Indonesia [31]. Despite the fact that the IDHS did not provide data on consumption expenditure, the DHS data represented the Indonesian population. Using the AF method and datasets from IDHS 2007, they concluded that child mortality, a health indicator with a tendency to be high, accounts for half of MPI in Indonesia. Moreover, IDHS 2012 was used by Alkire and Santos (2014) to examine further research. They also discovered that in both urban and rural areas of Indonesia, the length of schooling and child mortality were the main contributors to overall poverty. In terms of areas, data on urban child mortality, particularly in Jakarta, led to a higher deprivation measure than rural areas. Additionally, using SUSENAS data from 2011 to 2013 and the Alkire-Foster (AF) method, Indriani (2019) conducted a multidimensional poverty analysis in Central Java Province. The study measured a complex and multidimensional poverty phenomenon using three dimensions, such as nutrition and health, education, and living standards, and ten indicators, including household calorie consumption, household protein consumption, length of schooling, school participation, cooking fuel, sanitation, clean water, access to electricity, type of floor, and assets owned.

The results of the study demonstrated that poverty measured only in terms of the monetary dimension provides different insights compared to poverty measured in terms of multiple dimensions (multidimensional). Despite both projections pointing to the same phenomenon—that is, that the poverty rate in rural areas is higher than in urban areas—the fact that the multidimensional poverty percentage is higher than the monetary poverty percentage indicates that some people who were not considered to be poor monetarily were nonetheless poor when measured by multiple dimensions of poverty [33]. According to the AF method (Alkire and Foster, 2011) of measuring multidimensional poverty, a household is defined to be poor if they are deprived in several combination indicators and experience more than 33% of all deprivations. It indicates that a household must be deprived in at least two dimensions—either health and education, all standard of living indicators, or one of health or education and three standard of living indicators— in order to be classified as multidimensionally poor. Furthermore, based on the 33% cutoff line, a household that lives without electricity, lacks access to drinking water, and does not have an appropriate floor is not categorized as multidimensionally poor [37]. In addition, Martinetti (2006) made note of the fact that in developed and less developed regions, the length of formal education or schooling could reflect different levels of educational achievement.

This study measured poverty using the fuzzy C-means clustering (FCM) approach instead of considering it as a subject determined by the position of the household or individual in the distribution of some aspects of indicators. By clustering the data on poverty, this approach measures individuals' poverty levels to a certain degree rather than merely categorizing them as poor or not. A data classification technique referred to as clustering enables the determination of priority regions for the distribution of programs aimed at alleviating poverty (Betti et al., 2011; Pipino, 1981). Furthermore, clustering, a commonly used technique, is used to cluster data or objects into groups of data (clusters), in which each cluster consists of similar data and differs from the data in other clusters [40]. In several studies, such as those by Aassve et al. (2007), Belhadj (2011), Betti et al. (2011, 2015a, 2016), Betti and Verma (2008), and Verma et al. (2017), multidimensional indicators have been analyzed using FCM. Additionally, Verma et al. (2017) have proposed a new fuzzy measure of both monetary and non-monetary deprivation. In the case of non-monetary deprivation, explanatory and confirmatory factor analyses were employed to determine the dimensions or groups of initial items of deprivation, and a weighting scheme was then used to aggregate the individual items into the dimension they represented. The technique is applied to the nations of the European Union using data from the 2011 European Union-Statistics on Income and Living Conditions (EU-SILC) survey.

There are several cluster analysis-related studies in Indonesia. Kurniawan and Fatulloh (2017) employed the K-means algorithm to cluster social conditions in Batam, Indonesia, on the basis of social indicators such as poverty data, divorce data, the percentage of the population with the highest levels of high school education, population distribution by age, and morbidity. These indicators led to the development of four different clusters and characteristics. Moreover, Febrianti and Buana (2018) clustered the regencies and municipalities in East Java into four groups based on their analysis of the poverty indicators, which included eight indicators. Additionally, Risqiyani and Kesumawati (2016) employed the fuzzy C-means clustering (FCM) method to cluster regencies or municipalities in Central Java. The study took into account both the population factor and the health facilities and personnel factor due to the lack of health facilities, doctors, village midwives, the number of poor, the average household members, the population density level, the average educational level, and the human development index. The study generated four clusters, with cluster two containing 13 regencies and municipalities and cluster four containing 10 regencies and municipalities. Meanwhile, clusters one and three contain 7 and 13 regencies and municipalities, respectively.

Hidayat et al. (2017)used the K-means and the FCM method to cluster Central Java's regencies and municipalities based on factors affecting poverty in 2016. Four clusters were identified after employing the FCM method to cluster data. Among the four clusters, cluster one contained 18 regencies, cluster two contained 4 regencies and 1 municipality, cluster three contained 5 municipalities and 3 regencies, and cluster four contained 4 regencies. Furthermore, Mustafidah (2017) used SUSENAS data to cluster 35 regencies or municipalities in Central Java. The best clustering results, which were five clusters, were obtained employing three independent variables in the FCM method. Additionally, there were characteristics for each cluster, such as cluster one consisting of a high source of lighting, cluster two consisting of the type of roof used, cluster three consisting of a low lighting source and type of roof, cluster four consisting of being influenced by the average indicator per capita on food commodities, and cluster five consisting of the average expenditure on food commodities. Moreover, using the FCM algorithm, Nidyashofa and Istiawan (2017) clustered regencies or municipalities in Central Java based on the prosperity level in 2015. According to the study, Central Java could be categorized into three clusters of regencies or municipalities. The categorizations are as follows: Cluster one, which consisted of seven regencies or municipalities, was categorized as very poor; cluster two, which consisted of 11 regencies or municipalities, was categorized as poor; and cluster three, which consisted of 17 regencies or municipalities, was categorized as almost poor.

Previous studies have revealed that by comparing the degrees of poverty in different regencies or municipalities within a cluster, it was possible to investigate the leading causes of poverty in rural areas. Moreover, it was possible to group each member into clusters with various levels of membership in each of the areas, which led to the FCM method being employed. In general, the FCM algorithm could offer a more effective way to measure poverty in Central Java. Additionally, the main objectives of this study were to cluster multidimensional poverty and analyze poverty characteristics that differ at the regency level. Furthermore, Central Java Province was the subject of this study because it is the most populated province and where the majority of poor people reside. Given the enormous inter-island disparity, it was possible that the fact that the samples selected were in Java could bias the results. In addition, in Central Java, the poverty rate was higher in rural than urban areas, and those who live in poverty in rural areas tend to have low incomes and low purchasing power, poor nutrition, illiteracy, a high risk of infant and maternal mortality, and lower housing standards than those who live in poverty in urban areas (Quibria, 2002). The significance of this study is that a clustering method was used together with multidimensional poverty data in studying poverty in Central Java. Some pattern-analytical exploration, clustering, and decision-making may benefit from the use of cluster analysis (Martinetti, 2006). Moreover, in order to make policies and development strategies more precisely targeted and successful, poverty grouping is one method to determine the characteristics that define the level of people's prosperity in each region (Betti & Verma, 2008).

METHODOLOGY

This study primarily included data from regencies or municipalities in Central Java Province, consisting of 30 regencies, 5 municipalities, and 29,190 respondents. The dataset, which included 16 variables on the aspect of multidimensional poverty and covered 35 regencies or municipalities, is shown in Table 1.

Table 1: List of Variables

Code	Variables	Description
X_1	Non-food expenditure per capita	Percentage of the population with sufficient non-food consumption (frequently purchased goods and services, such as soap and transport expenses), as well as less often yet routinely purchased items, such as clothing and kitchen equipment.
X_2	Food expenditure per capita	Percentage of the population with sufficient foods, including food consumed by households from all attainable sources, such as purchases (with cash or barter), consumption of food produced at home, and transfers made in the form of gifts or in-kind payments.
X_3	Population density	Percentage of the population or concentration of individuals within a specific geographic locale.
X_4	Open unemployment rate	Percentage of the population of working age who are unemployed, are available for work, and have made efforts to obtain employment.
X_5	Drinking water adequacy	Percentage of the population who have access to sufficient water to drink.
X_6	Source of water	Percentage of the population using branded bottled water, refilled water, taps, boreholes or pumps, and protected wells as sources of drinking water.
X_7	Access to sanitation facilities	Percentage of the population with septic tank or WWTP sewerage.
X_8 X_9	7–24 years old, never studied (no longer studying)	Percentage of the population aged 7–24 years old who never attended school or are no longer studying.
X_{10}	*PKH* recipient	Percentage of the population currently receiving or having received PKH and KKS programs.
X_{11}	Literacy rate	Percentage of the population who are able to read 1 or more: latin/alphabet letters, Arabic/*hijaiyah* letters, other letters.
X_{12}	Educational level	Percentage of household heads with a minimum education of junior high school or equivalent (9 years of compulsory education).
X_{13} X_{14}	Ownership	Percentage of the population with 1 or more: car and television 30".

X_{15}	*KIP* recipient	Percentage of the population currently receiving or having received *PKH* and *KKS* programs.
X_{16}	Food assistance	Percentage of the population having or currently receiving non-cash or food assistance.

Clustering is a multivariate statistical method to cluster objects based on proximity or resemblance measures. It differs from grouping, which requires all members of a group to have the same condition. Meanwhile, the proximity of available sample characteristics (average proximity) is the basis for clustering, one of which uses the Euclidean distance. The fuzzy clustering method implements its members' fuzziness as a weighting basis for clustering [53]. The essential advantage of the fuzzy clustering method is that it may produce grouping results for objects that are erratically spread. Moreover, it is crucial to consider the tendency of data points to cluster for the reason that if there is an erratic distribution of data, it is possible that a data point has properties or characteristics similar to those of other clusters. Additionally, fuzzy C-means (FCM), which had been developed by Bezdek (1981) and initially proposed by Dunn (1973), is one of the options available for the data clustering algorithm. The fuzzy c-means is a clustering method developed from c-means by applying its members' fuzziness. Identifying the cluster center—which will serve as the average position of each cluster—is the key concept of FCM. Each data point has a different level of membership in each cluster. A curve representing the membership function demonstrates how input data points are mapped into membership values. In the FCM method, the membership function variable refers to how probable it is for data to become a member of a cluster.

RESULTS AND DISCUSSION

In order to generate the initial set of clusters and fuzzy membership degree matrix, FCM and other alternating optimization algorithms must be run as an initialization phase. Moreover, to determine the optimal number of clusters, FCM algorithms require clusters to be pre-specified. The number of clusters in the investigated dataset has been determined using a variety of methods. In this study, the elbow method, one of the cluster analysis methods, was employed to determine the optimal number of clusters. The elbow method is one of the methods used to determine the optimal number of clusters. This method was implemented using the R programming language because it offers a set of software for data calculation, simulation, and graphics display, as well as an interpreted programming language.

The elbow is the first method to distinguish the optimal number of clusters for the dataset under analysis. The elbow method is based on a visual interpretation of a graphical plot of the number of clusters (k) by examining the percentage comparison of the number of clusters at an angle at a point. The elbow method results in a slightly subjective estimate of the potential optimal number of clusters for the dataset under study. The main concept is to specify k=2 as the initial optimal number of clusters k, then repeatedly increase k by intervals of one to the estimated optimal number of clusters, and finally distinguish the optimal number of clusters k that corresponds to the plateau (Cattell, 1966).

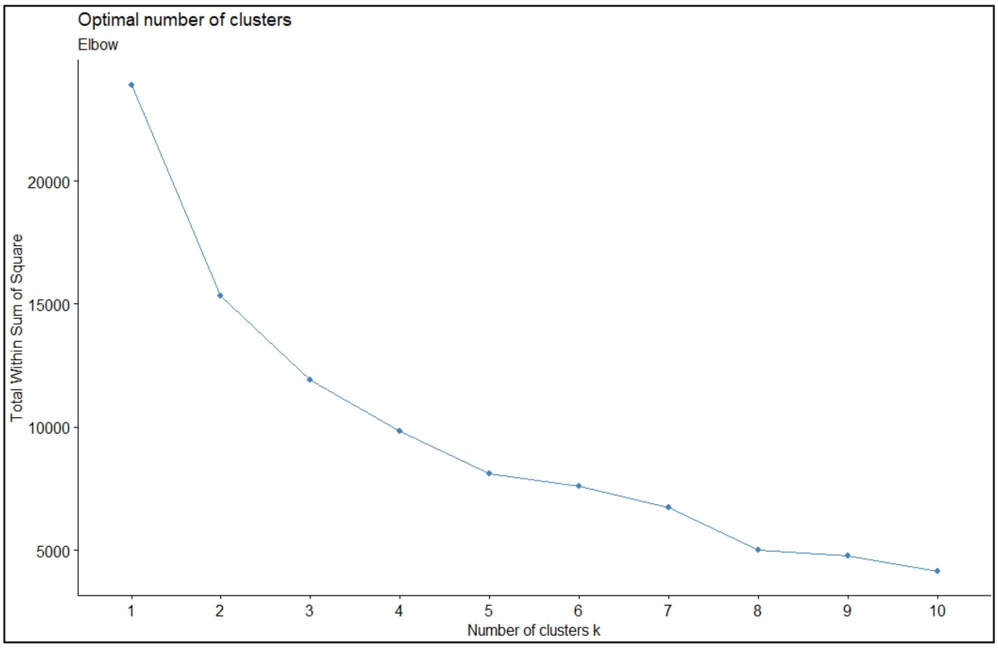

Figure 1. Elbow Graph

An optimal scree plot for the elbow method is one that begins steeply and bends at an elbow. The curve becomes flatter after the elbow. As stated by Cattell (1966), the elbow is used as the cutting-off point. The elbow-shaped sum of the square value, which has significantly decreased, is used to estimate the best cluster value. In other words, the sum of the square value decreases as the number of clusters k increases. The sum of the squares is the total distance from a centroid to data in the same cluster. Figure 1 illustrates that the X-axis represents the k value, and the Y-axis represents the sum of the square values. In addition, according to Figure 1, the optimal clusters were estimated to be 2 or 4 clusters since these clusters provide angles that form an angle at the point of greatest decrease.

The optimal number of clusters is 2 or 4 clusters, as demonstrated in the previous explanation. Validation is required by conducting tests on several clusters before the FCM machine learning model is run. Cluster validation evaluates how well the clustering results are done. There are sets of suitable sizes for each cluster that is formed, such as the cluster validity index value. According to Brock et al. (2008) and Charrad et al. (2014), there are three types of cluster validation techniques: internal, external, and relative. Without referring to external data, internal cluster validation techniques evaluate the quality of cluster results using internal data from the clustering process. In addition, internal cluster validation techniques were employed in this study.

In this test, these measurement parameters were used, including the Dunn index, the Daviex Bouldin (DB) index, the Xien Beni (XB) index, and the Modified Partition Coefficient (MPC), respectively. According to Brock et al. (2008), the Dunn index is the ratio value between the closest distance between observation data from distinct clusters and the greatest intra-cluster distance. Furthermore, one of the most used methods for cluster validation is the Xien-Beni index method (XB), which was initially developed by Xie and Beni. The proximity of the data within a cluster and the distance between cluster centers are the main focus of the XB index method. The optimal number of clusters is signified by the smallest XB value (Mota et al., 2017). In addition to the Dunn and XB indexes, the Daviex Bouldin (DB) index is a method for validating clustering. The DB index value that approaches 0 implies the better the cluster obtained (Pakhira et al., 2004). The primary objective of this validation is to ensure that, when compared to other clusters, the cluster with four groups is the best.

Table 2: Cluster Validation

n Cluster	***Dunn***	***Davies Bouldin* (DB)**	***Xien Beni* (XB))**	***Modified Partition Coefficient* (MPC)**
2	0.2136	1.2623	0.4797	0.2399
3	0.3069	1.3524	0.7865	0.2624
4	0.3658	1.1164	0.4077	0.2922
5	0.2384	1.1776	0.6368	0.2708

Table 2 shows the validity index value for the FCM method using 2, 3, 4, and 5 clusters. Additionally, the maximum Modified Partition Coefficient (MPC), Dunn indices, and the minimum Davies Bouldin (DB) and Xie Beni (XB) indices were used to determine the optimal number of clusters. Furthermore, Table 2 shows the maximum validity index values of Duun and MPC for 4 clusters, with an index value of 0.3568 for the Dunn index and 0.2928 for the MPC index. Table 2 also shows the minimum DB and XB validity index values for 4 clusters, with an index value of 1.1164 for the DB index and 0.4077 for the XB index. In addition, the optimal number of clusters was determined to be four based on the validity index value obtained.

Following the determination of the number of clusters to be run, the next step was to determine the members of each cluster. Fuzzy clustering was applied to a subset of data consisting of 29 regencies and 6 municipalities in Central Java, and the data were scaled to convert into a data frame, generating four clusters as the desired number of clusters. Moreover, the membership degrees of the nearby clusters were associated with the data points in the cluster's center, which can have a degree equal to 1. Additionally, Figure 2 is a convex-type cluster plot that aims to find the best membership for each object (Rousseeuw, 1995). Furthermore, clusters 1, 2, 3, and 4 each had a certain point close to the cluster's boundary that was recognized. In Figure 2, four distinct colors—blue, yellow, gray, and red—were used to represent the clusters 1, 2, 3, and 4, respectively. The details of the regency or municipality's codes and names in Figure 2 can also be seen in Appendix 2.

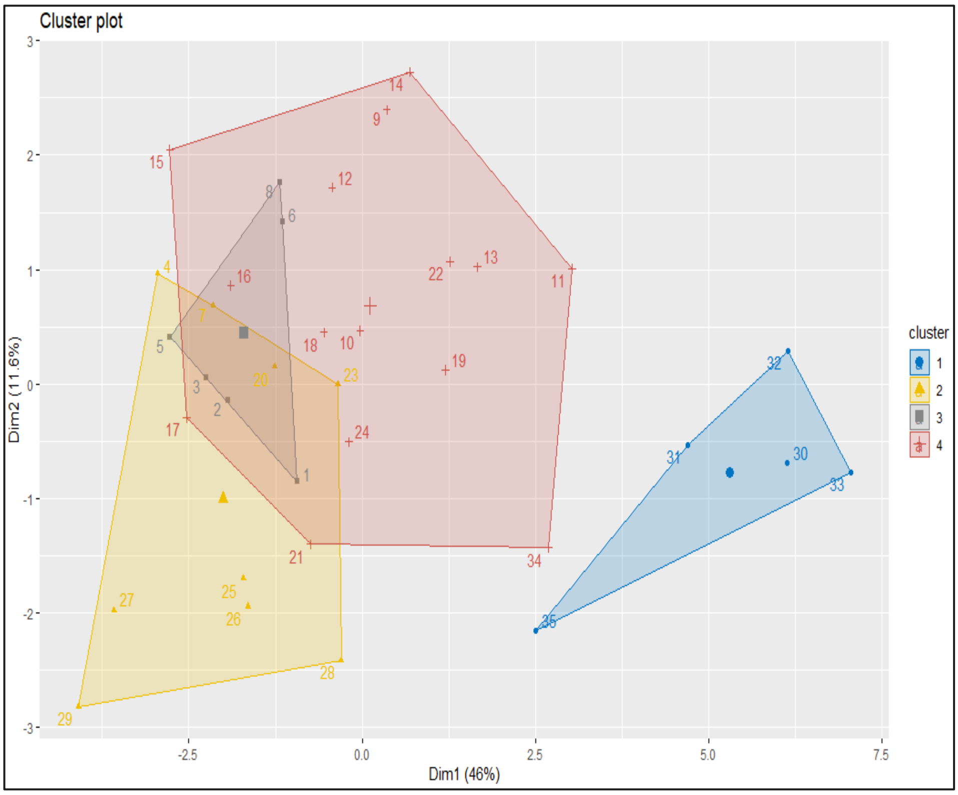

Figure 2: Cluster Plot

Table 3: Regencies or Municipalities of Each Cluster

Cluster	Regency or Municipality
1	Magelang Municipality, Surakarta Municipality, Salatiga Municipality, Semarang Municipality, and Tegal Municipality
2	Banjarnegara, Wonosobo, Jepara, Temanggung, Batang, Pekalongan, Pemalang, Tegal, Brebes
3	Cilacap, Banyumas, Purbalingga, Kebumen, Purworejo, Magelang
4	Boyolali, Klaten, Sukoharjo, Wonogiri, Karanganyar, Sragen, Grobongan, Blora, Rembang, Pati, Kudus, Demak, Semarang, Kendal, Pekalongan Municipality.

Figure 2 explains the four clusters with an eigenvalue of 46%, which means the cluster plot explains 46% of the variables used. Figure 2 also shows the clustering for the mapping of 29 regencies and 6 municipalities in Central Java based on the poverty level. Regencies or municipalities were grouped into four clusters: cluster 1 contains five municipalities; cluster 2 contains nine regencies; cluster 3 contains six regencies; and cluster 4 contains 14 regencies and 1 municipality. Table 3 shows the details of the regencies or municipalities of each cluster that have been formed using FCM clustering, and Figure 3 displays the map of multidimensional poverty resulting from cluster results.

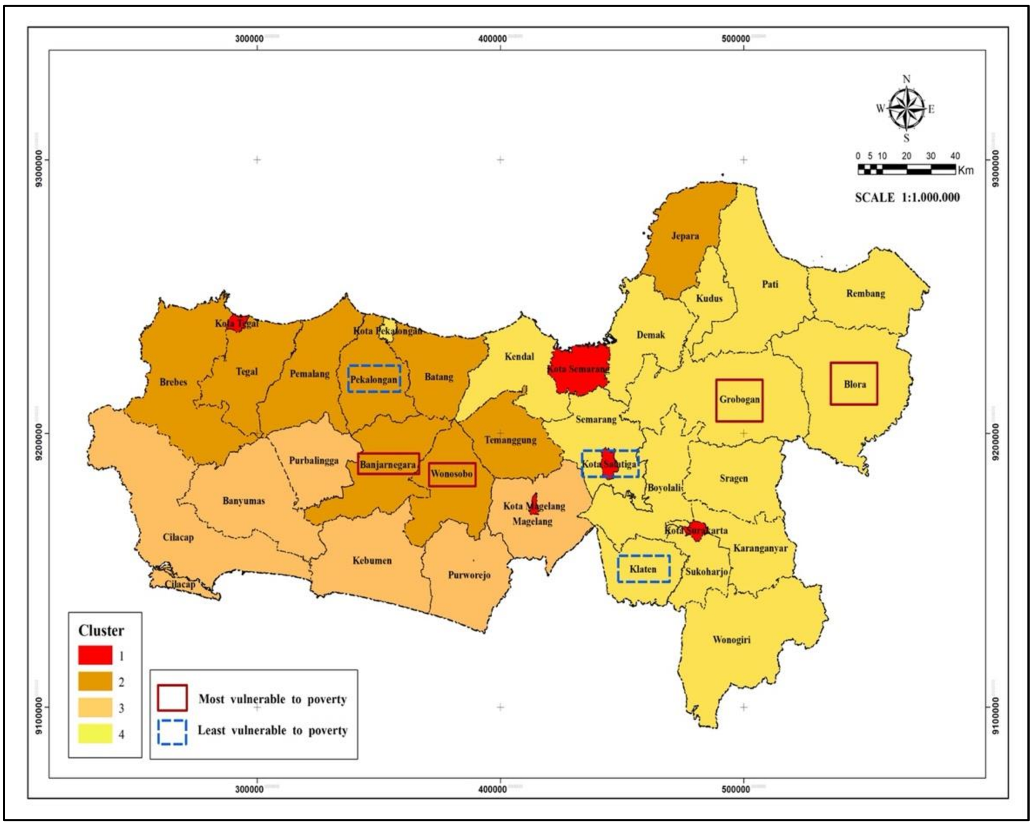

Figure 3: Regencies or Municipalities by Cluster

Table 4 shows the fuzzy C-means clustering (FCM) method was used to cluster regencies or municipalities in Central Java based on the level of poverty, with the optimal number of clusters being 4 clusters. The centroid of each characteristic served as the basis for each cluster's ranking. Moreover, the most significant average for the multidimensional poverty variables provides insight into which variables received the highest ranking. For instance, the average value for the percentage of non-food expenditure per capita for cluster 1 was 58.71. Therefore, cluster 1 was ranked fourth, indicating that cluster 1 was the least poor within this variable. In general, based on the average value of the optimal cluster results, it can be inferred that cluster 2 was the cluster with the highest rank in the poor category, which means they were most likely to be living in poverty, and cluster 4 was ranked second. Furthermore, cluster 3 was ranked third, and cluster 1 was ranked fourth. In addition, the municipalities of Central Java in cluster 1 had the lowest multidimensionally poor ranking with around 15% (5 of 35 regencies and municipalities).

Table 4: Means of Variables of Each Cluster (%)

Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Non-food expenditure per capita	58.71	48.11	51.11	50.76
Food expenditure per capita	41.29	51.89	48.89	49.24
Population density	1.51	3.25	3.72	2.73
Open unemployment rate	8.38	6.99	5.93	5.57
Drinking water adequacy	99.46	93.08	90.97	95.98
Source of water	93.27	77.29	74.09	83.62
Access to sanitation facilities	91.13	67.20	81.60	91.13
7–24 years old, never studied	0.15	0.31	0.25	0.35
7–24 years old, no longer studying	26.20	33.75	28.06	30.37
PKH recipient	8.30	17.65	18.91	13.11
Literacy rate	97.52	92.50	94.26	92.44
Educational level (9 years or fewer)	46.55	77.18	71.69	66.95
Ownership (Car)	16.61	9.05	9.24	12.14
Ownership (TV)	22.36	10.02	9.47	10.29
KIP recipient	8.94	14.82	14.82	11.35
Food assistance	12.20	23.24	27.82	20.85
Number of regencies or municipalities	**5**	**9**	**6**	**15**

The municipalities in cluster 1 were spread over all urban areas in Central Java and were consistent with the poverty index from Statistics Indonesia (2020a), which reported that municipality areas have a lower poverty rate (below 10%). This result also highlights the uneven development progress and the impact of heavy industrial agglomerations concentrated in urban areas. In addition, based on the average of each variable, regions in cluster 1 had issues related to the open unemployment rate. It may be due to the tendency of people to migrate to big cities, and job limitation was also one of a variety of reasons to improve their living standards (Obermayr, 2017). Meanwhile, cluster 2 was the cluster that had the most multidimensional poverty problems compared to the other clusters. Based on the means of each variable, the regencies included in cluster 2 have issues with the source of water, access to sanitation facilities, 7-24 years old no longer studying, and educational level. This region also included the population density and the open unemployment rate items. Additionally, it is important to pay attention to education indicators because there was a significant percentage of the population with the highest education at a maximum of junior high school (77.1%). This suggests that the majority of the population in cluster 2 only completed junior high school or below. According to the perspective of education, the lower a person's education, the more vulnerable they are (Ahluwalia & Montek, 1985; Bhagwati, 1988). This is because having a lower level of education limits people's access to and capacity to obtain employment. This is not only limited to formal and non-formal education, such as training.

In cluster 2, there is still limited access to sanitation and a source of water for cooking or washing. Geographically, cluster 2 included areas that are mountainous and/or seaside, making them more vulnerable to having limited access to sanitation and sources of water. The findings are in line with the study by Quibria (2002), who stated that a number of factors, including regional conditions, had an impact on the poverty rate and that the impact of poverty in regions was higher than in cities. The clustering's results are consistent with the results of Statistics Indonesia (2021). Moreover, the results of this study are in line with research conducted in Indonesia by Hanandita and Tampubolon (2016) and Indriani (2019), who asserted that Wonosobo and Banjarnegara Regencies had the highest poverty rate and were the most multidimensionally poor. Moreover, the results remained the same using different indicators, such as the number of illnesses suffered, length of illness suffered, and literacy rate. Despite the fact that the indicators were different, the finding further confirms that Banjarnegara and Wonosobo Regencies were areas with a high multidimensional poverty rate. Moreover, cluster 3 is a cluster that has poverty issues related to population density, open unemployment rate, and source of water. The six regencies included in cluster 3 were spread over most of southern Central Java. In contrast, cluster 4 includes the majority of Central Java's regions and has issues with food expenditure, 7-24 years old never studied, 7-24 years no longer studying, and literacy rate. Due to the fact that cluster 4 includes around half of the regions in Central Java, education level (66.9%) is the dominant issue. This indicates that only 33.1% of the 14 regions in cluster 4—or nearly half of the total regencies in Central Java—have residents who have completed more than nine years of compulsory education.

CONCLUSION

The average poverty rate in Central Java was 14.2%, which was still relatively high and was classified as hard core (>10%), indicating that poverty reduction policies implemented by the government have not run optimally. In other words, the government has not been able to achieve the National Medium-Term Development Plan of decreasing the number of people in absolute poverty to 8–10% of the population, improving income distribution with family-based social protection, expanding economic opportunities for low-income people, and fulfilling their basic needs. Moreover, this study discovered that the optimal number of clusters was four using fuzzy C-means clustering. Furthermore, the poorest regions with a low education level, non-food expenditure per capita, access to sanitation facilities, a high population density, and a high open unemployment rate were found in cluster 2, which was followed by cluster 3 and cluster 4. Meanwhile, cluster 1 includes the regions with the lowest levels of poverty in urban areas. However, half of the regions in Central Java were in cluster 4, indicating that cluster 4 contained a significant number of mostly poor people.

In this study, the majority of the clusters demonstrated a low education level, a high population density, a high open unemployment rate, and a high rate of 7-24 years old no longer studying. This result shows that unemployment is directly associated with low levels of education. It also confirms that education is still a substantial problem in Central Java. Lack of education limits a person’s ability to find opportunities. In addition, clusters with lower education levels and employment rates have a tendency to have less adequate drinking water and access to sanitation. Furthermore, the result of the cluster analysis revealed that an important characteristic of poverty in Central Java is the significant variation dependent on geographical location. The prevalence of poverty was often higher in the western and southern regencies or municipalities than it was in the eastern and northern regencies or municipalities. This finding is in line with the statement of Quibria (2002) which regencies tend to have a higher poverty rate than cities area due to geographical location.

According to the results of this study, 5 out of 16 indicators used to measure the poverty rate in Central Java were statistically significant. The significant indicators are non-food expenditure, drinking water adequacy, access to sanitation facilities, the open unemployment rate, and ownership (TV). This study also revealed a large gap in the open unemployed rate variable and access to sanitation facilities among regencies and municipalities in Central Java. Furthermore, urban areas (cluster 1) with an unemployment problem have a specific pattern of cluster formation. Meanwhile, in rural areas (clusters 2, 3, and 4), the significant indicators are non-food expenditure, source of water, access to sanitation facilities, and education level. From the clustering results, it can be concluded that the issues faced by each regency and municipality in Central Java are different; employment issues dominate urban areas, and education and sanitation issues dominate rural areas. In addition, these results may provide insight for the government to set priorities for regional development in each area.

REFFERENCES

Aassve, A., Betti, G., Mazzuco, S., & Mencarini, L. (2007). Marital disruption and economic well-being: a comparative analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *170*(3), 781–799. https://doi.org/10.1111/j.1467-985X.2007.00483.x

Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, *95*(7–8), 476–487. https://doi.org/10.1016/j.jpubeco.2010.11.006

Alkire, S., Kangaratnam, U., & Suppa, N. (2019). *The Global Multidimensional Poverty Index (MPI) 2019* (47).

Alkire, S., Oldiges, C., & Kanagaratnam, U. (2021). Examining multidimensional poverty reduction in India 2005/6–2015/16: Insights and oversights of the headcount ratio. *World Development*, *142*, 105454. https://doi.org/10.1016/j.worlddev.2021.105454

Alkire, S., & Santos, M. E. (2014). Measuring Acute Poverty in the Developing World: Robustness and Scope of the Multidimensional Poverty Index. *World Development*, *59*, 251–274. https://doi.org/10.1016/j.worlddev.2014.01.026

Ballon, P., & Apablaza, M. (2012). Multidimensional Poverty Dynamics in Indonesia. *Paper Presented at the Research Workshop on Dynamic Comparison between Multidimensional Poverty and Monetary Poverty.*

BAPPENAS. (2007). *Law of The Republic of Indonesia Number 17 of 2007 on Long-Term National Development Plan of 2005-2025* (Economic and Social Policy).

Battiston, D., Cruces, G., Lopez-Calva, L. F., Lugo, M. A., & Santos, M. E. (2013). Income and Beyond: Multidimensional Poverty in Six Latin American Countries. *Social Indicators Research*, *112*(2), 291–314. https://doi.org/10.1007/s11205-013-0249-3

Baulch, B., & Masset, E. (2003). Do Monetary and Nonmonetary Indicators Tell the Same Story About Chronic Poverty? A Study of Vietnam in the 1990s. *World Development*, *31*(3), 441–453. https://EconPapers.repec.org/RePEc:eee:wdevel:v:31:y:2003:i:3:p:441-453

Belhadj, B. (2011). New fuzzy indices of poverty by distinguishing three levels of poverty. *Research in Economics*, *65*(3), 221–231. https://doi.org/10.1016/j.rie.2010.10.002

Bérenger, V. (2017). Using ordinal variables to measure multidimensional poverty in Egypt and Jordan. *The Journal of Economic Inequality*, *15*(2), 143–173. https://doi.org/10.1007/s10888-017-9349-7

Betti, G., D'Agostino, A., & Neri, L. (2011). Educational Mismatch of Graduates: a Multidimensional and Fuzzy Indicator. *Social Indicators Research*, *103*(3), 465–480. https://doi.org/10.1007/s11205-010-9712-6

Betti, G., Gagliardi, F., Lemmi, A., & Verma, V. (2015). Comparative measures of multidimensional deprivation in the European Union. *Empirical Economics*, *49*(3), 1071–1100. https://doi.org/10.1007/s00181-014-0904-9

Betti, G., Soldi, R., & Talev, I. (2016). Fuzzy Multidimensional Indicators of Quality of Life: The Empirical Case of Macedonia. *Social Indicators Research*, *127*(1), 39–53. https://doi.org/10.1007/s11205-015-0965-y

Betti, G., & Verma, V. (2008). Fuzzy measures of the incidence of relative poverty and deprivation: a multi-dimensional perspective. *Statistical Methods and Applications*, *17*(2), 225–250. https://doi.org/10.1007/s10260-007-0062-8

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Springer US. https://doi.org/10.1007/978-1-4757-0450-1

Bossert, W., D'Ambrosio, C., & Peragine, V. (2007). Deprivation and Social Exclusion. *Economica*, *74*(296), 777–803. http://www.jstor.org/stable/4541571

Castaneda Aguilar, R. A., Doan, D. T. T., Newhouse, D. L., Nguyen, M. C., Uematsu, H., & Azevedo, J. P. W. de. (2016). *Who are the poor in the developing world?* (7844; Policy Research Working Paper Series).

Chakravarty, S. R. (2006). *An axiomatic approach to multidimensional poverty measurement via fuzzy sets', in A. Lemmi and G. Betti (Eds.), Fuzzy Set Approach to Multidimensional Poverty Measurement*.

Chakravarty, S. R., & D'Ambrosio, C. (2006). THE MEASUREMENT OF SOCIAL EXCLUSION. *Review of Income and Wealth*, *52*(3), 377–398. https://doi.org/10.1111/j.1475-4991.2006.00195.x de Muro, P., Mazziotta, M., & Pareto, A. (2011). Composite Indices of Development and Poverty: An Application to MDGs. *Social Indicators Research*, *104*(1), 1–18. https://doi.org/10.1007/s11205-010-9727-z

Duclos, J., Sahn, D. E., & Younger, S. D. (2006). Robust Multidimensional Poverty Comparisons. *The Economic Journal*, *116*(514), 943–968. https://doi.org/10.1111/j.1468-0297.2006.01118.x

Dunn, J. C. (1973). A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics*, *3*(3), 32–57. https://doi.org/10.1080/01969727308546046

Febrianti, Fitria. A., & Buana, A. (2018). K-Means Clustering Dengan Metode Elbow Untuk Pengelompokan Kabupaten Dan Kota Di Jawa Timur Berdasarkan Indikator Kemiskinan. *Seminar Nasional Hasil Riset Dan Pengabdian* .

Foster, J., Greer, J., & Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica*, *52*(3), 761. https://doi.org/10.2307/1913475

Hanandita, W., & Tampubolon, G. (2016). Multidimensional Poverty in Indonesia: Trend Over the Last Decade (2003–2013). *Social Indicators Research*, *128*(2), 559–587. https://doi.org/10.1007/s11205-015-1044-0

Hidayat, R., Wasono, R., & Darsyah, M. Y. (2017). Pengelompokkan Kabupaten/Kota Di Jawa Tengah Menggunakan Metode K-Means dan Fuzzy C-Means. *Seminar Nasional Pendidikan, Sains Dan Teknologi Fakultas Matematika Dan Ilmu Pengetahuan Alam*, 240–250.

Indriani, L. (2019). Analisis Kemiskinan Multidimensi di Provinsi Jawa Tengah, 2011-2013. *Jurnal Aplikasi Statistika & Komputasi Statistik*, *10*(2), 13. https://doi.org/10.34123/jurnalasks.v10i2.72

Kurniawan, Ely. D., & Fatulloh, A. (2017, February 12). Clustering of Social Conditions in Batam, Indonesia Using K-Means Algorithm and Geographic Information System. *International Conference on Earth Sciences and Engineering* .

Martinetti, E. C. (2006). Capability Approach and Fuzzy Set Theory: Description, Aggregation and Inference Issues. In A. Lemmi & G. Betti (Eds.), *Fuzzy Set Approach to Multidimensional Poverty Measurement* (pp. 93–113). Springer US. https://doi.org/10.1007/978-0-387-34251-1_6

McKay, A., & Lawson, D. (2003). Assessing the Extent and Nature of Chronic Poverty in Low Income Countries: Issues and Evidence. *World Development*, *31*(3), 425–439. https://EconPapers.repec.org/RePEc:eee:wdevel:v:31:y:2003:i:3:p:425-439

Ministry for Human Development and Culture. (2020, February 11). *Menko PMK akan Koordinasikan Program Penurunan Tingkat Kemiskinan dan Tingkat Stuntin*. Kesehatan.

Ministry of Social Affairs. (2021). *Pedoman Pelaksanaan Program Keluarga Harapan* .

Mustafidah, R. (2017). *Pengelompokan Kabupaten/Kota Di Provinsi Jawa Tengah Berdasarkan Indikator Kemiskinan Dengan C-Means dan Fuzzy C-Means*. Institute Teknologi Sepuluh November.

Nidyashofa, N., & Istiawan, D. (2017). Penerapan Algoritma Fuzzy C-Means untuk Pengelompokan Kabupaten/Kota di Jawa Tengah Berdasarkan Status Kesejahterann Tahun 2015. *The 6thUniversity Research Colloquium 2017*, 23–30.

OPHI. (2011). *Country Briefing: Indonesia Multidimensional Poverty Index (MPI)* (MultidimensionalPoverty Index Data Bank, Oxford Poverty and Human Development Initiative).

OPHI. (2019). *Global MPI Country Briefing 2019: Indonesia (East Asia and the Pacific)* (MultidimensionalPoverty Index Data Bank, Oxford Poverty and Human Development Initiative).

Pipino, L. Lg. J. P. van. (1981). POTENTIAL IMPACT OF FUZZY SETS ON THE SOCIAL SCIENCES. *Cybernetics and Systems*, *12*(1–2), 21–35. https://doi.org/10.1080/01969728108927662

Quibria, M. G. (2002). *Growth and Poverty: Lessons from the East Asian Miracle Revisited* (33; Research Paper Series).

RAND. (2007). *IFLS-4 Survey Description*.

Ravallion, M., Chen, S., & Sangraula, P. (2009). Dollar a Day Revisited. *The World Bank Economic Review*, *23*(2), 163–184. https://doi.org/10.1093/wber/lhp007

Risqiyani, T. A., & Kesumawati, A. (2016). Pengelompokan Kabupaten Kota di Provinsi Jawa Tengah dengan Fuzzy C-Means Clustering (Studi Kasus : Jumlah Kasus Gizi Buruk, Faktor Sarana dan Tenaga Kesehatan serta Faktor Kependudukan di Jawa Tengah Tahun 2014). *Seminar Nasional Matematika Dan Pendidikan Matematika UNY 2016*.

Sahn, E. , D., & Stifel, D. (2003). Exploring Alternative Measures of Welfare in the Absence of Expenditure Data. *Review of Income and Wealth*, *49*(4).

Santos, M. E., Dabus, C., & Delbianco, F. (2019). Growth and Poverty Revisited from a Multidimensional Perspective. *The Journal of Development Studies*, *55*(2), 260–277. https://doi.org/10.1080/00220388.2017.1393520

Scott, L. (2002). A poverty indicator system for local government. *Development Southern Africa*, *19*(4), 483–501. https://doi.org/10.1080/0376835022000019437

Sen, A. (1976). Poverty: An Ordinal Approach to Measurement. *Econometrica*, *44*(2), 219. https://doi.org/10.2307/1912718

Statistics Indonesia. (2020). *Profil Kemiskinan di Indonesia Maret 2020*.

Statistics Indonesia. (2021). *Persentase Penduduk Miskin (semesteran) (Persen), 2010-2021*.

Thomas, D., Witoelar, F., Frankenberg, E., Sikoki, B., Strauss, J., Sumantri, C., & Suriastini, W. (2012). Cutting the costs of attrition: Results from the Indonesia Family Life Survey. *Journal of Development Economics*, *98*(1), 108–123. https://doi.org/10.1016/j.jdeveco.2010.08.015

United Nations. (2015). *The Millenium Development Goals Report*.

United Nations. (2018). *Multidimensional Poverty in Comoros* (18–00003; Economic and Social Commission for Wester and Asia (ESCWA)).

Verma, V., Betti, G., & Gagliardi, F. (2017). Fuzzy Measures of Longitudinal Poverty in a Comparative Perspective. *Social Indicators Research*, *130*(2), 435–454. https://doi.org/10.1007/s11205-015-1194-0

Whelan, C. T. (2004). Understanding the Mismatch Between Income Poverty and Deprivation: A Dynamic Comparative Analysis. *European Sociological Review*, *20*(4), 287–302. https://doi.org/10.1093/esr/jch029

World Bank. (2017). *Monitoring Global Poverty: Report of the Commission on Global Poverty*.

Yu, J. (2013). Multidimensional Poverty in China: Findings Based on the CHNS. *Social Indicators Research*, *112*(2), 315–336. https://doi.org/10.1007/s11205-013-0250-x