Mapping model for target achievement of village SDGs using the Ensemble ROCK method: A case study of Sidoarjo Regency, East Java

Agnes Tuti Rumiati ^{(a)*} Nadhifa Zulfa Salsabila ^(b) Annisa Raina Nabilah T ^(c)
Henni Jovita Sari ^(d) Lulus Faticatu Riza ^(e)

^(a,b,c,d,e) Department of Statistics, Faculty of Science and Data Analytics Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

ARTICLE INFO

Article history:

Received 28 December 2021

Received in rev. form 01 Feb. 2022

Accepted 06 February 2022

Keywords:

Village SDGs, Ensemble ROCK, Clustering, Sidoarjo, SDGs

JEL Classification:

B55, O35, P23

ABSTRACT

Sustainable Development Goals (SDGs) are a set of goals for sustainable development which are consist of 17 goals that are expected to be achieved by 2030. 91% of Indonesia's territory is a rural area, consisting of more than 17,000 islands, almost 7500 villages. The villages contribute about 74% to the achievement of the national SDGs. Indonesia implements The Village SDGs which have 18 goals. The 18th goal is to achieve dynamic village institutions and adaptive village culture designed to accommodate various local cultures that characterize village management and are still applied by the community. Almost all district in Indonesia do not yet have a mapping of the status of achieving the Village SDGs. It will be easier for the district government to prioritize village development programs according to the needs of each village. In this study, research variables that measure 17 village goals will be used and grouped into 4 aspects, namely economic aspects, environmental aspects, health and education aspects and welfare aspects. Because village SDGs data is a mixture of numeric and categorical, ROCK ensemble analysis is used which is a grouping method for mixed data. In this study, it was found that the majority of villages in Sidoarjo Regency were still lagging behind in the fields of economy, health, education, and community welfare. This can be seen in the number of villages that are included in the low group based on the results of the analysis.

© 2022 by the authors. Licensee SSBFNET, Istanbul, Turkey. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Indonesia became one of the countries that agreed to Sustainable Development Goals (SDGs) program, and it was stated in Presidential Regulation No. 59 of 2017 concerning the Implementation of Achieving Sustainable Development Goals. SDGs is a global development agenda guided by five principles that harmonize the concepts of economic, social, environmental and government, namely people, planet, prosperity, peace, and partnership. (UN, 2015). SDGs has 17 goals that are expected to be achieved by 2030 Indonesia became one of the countries that agreed to this and it was stated in Presidential Regulation No. 59 of 2017 concerning the Implementation of Achieving Sustainable Development Goals. The objective of the SDGs is explicitly to mobilize development that leads to progress, especially in efforts to end all forms of poverty and hunger, reduce inequality, preserve the environment, and take steps for climate change (Iskandar, 2020).

Indonesia consists of more than 17,000 islands, around 7500 villages so that regionally 91% of Indonesia's territory is rural area. Therefore, villages contribute 74% to the achievement of the Sustainable Development Goals (Iskandar, 2020) with 11 sustainable national development goals that are closely related to village territory. Meanwhile, if viewed from the aspect of citizenship, there are

* Corresponding author. ORCID ID: 0000-0001-6551-1109

© 2022 by the authors. Hosting by SSBFNET. Peer review under responsibility of Center for Strategic Studies in Business and Finance.
<https://doi.org/10.20525/ijrbs.v11i1.1567>

Citation: Tuti Rumiati, A., Salsabila, N. Z., Nabilah T, A. R., Sari, H. J., & Riza, L. F. (2022). Mapping model for target achievement of village SDGs using the Ensemble ROCK method: A case study of Sidoarjo Regency, East Java. *International Journal of Research in Business and Social Science* (2147- 4478), 11(1).

43% of the Indonesian population in villages with 6 sustainable national development goals that are closely related to village residents.

The Village SDGs have 18 goals, whereas the 18th goal is to achieve dynamic village institutions and adaptive village culture designed to accommodate various local cultures that characterize village management and are still applied by the community. The implementation of the Village SDGs is aimed at directing development at the village level where to achieve the village SDGs with 18 objectives. The Village SDGs are divided into 8 goal groups which are then referred to as fields, those are poverty and hunger, health, education, women's friendliness, economy, environment, networked villages, and cultural responsiveness. The achievement of the Village SDGs is measured based on the Village SDGs indicators in accordance with the Regulation of the Minister of Villages 21 of 2020.

Analysis of the relationship between indicators and dimensions of sustainable development in Indonesia had been conducted by Setianingtias et al. (2019). Using the Structural Equation Modeling – Partial Equation Modeling (SEM-PLS) method, the results showed that there were 19 indicators that influenced sustainable development. The SEM-PLS modeling showed significant relationship between the dimensions of economy with other dimension (environment, social, institutional), but there was an insignificant relationship, namely the relationship between institutions and the environment.

Almost all districts in Indonesia do not yet have a mapping of the status of achieving the Village SDGs. With the mapping of villages in a district. Mapping model of the achievement of Village SDGs will be easier for the district government to prioritize village development programs according to the needs of each village.

This research takes a case study in Sidoarjo Regency, East Java province, Indonesia. Sidoarjo Regency consists of 18 sub-districts and 322 villages is a satellite city of the city of Surabaya. The Village SDGs data collection was developed by the Ministry of Villages and PDDT through the Village Information System (SID) and Development Village Index (IDM) application. In this study, the data used was IDM data which was selected according to the Village SDGs indicators. IDM is a composite index formed based on three indices, namely the social resilience index, economic resilience and ecological/environmental resilience (KDPDDT, 2021). IDM was developed based on the concept that to get to a developed and independent village, a sustainable development framework is needed where the aspects are social, economic and environmental.

Because the village SDGs data is a mixture of numeric and categorical, the ROCK ensemble analysis is used which is a grouping method for mixed data. Research related to the ROCK ensemble method has been carried out by several researchers, namely by (Ramdhany, 2018) For grouping villages in Bonodowoso Regency based on mixed numerical and categorical data in the case of the Village Development Index (IDM) by comparing two methods, namely the two-step cluster with the ensemble ROCK method, it was found that the best grouping was using the ensemble ROCK method. In addition, the ensemble method has also been used by Alvionita (2017) namely about the ensemble ROCK and SWFM methods in the case study of grouping citrus accessions and the results showed that the ROCK ensemble method gave better grouping results.

The results of this study will be useful as a basis for preparing program plans related to the achievement of the Village SDGs targets for Sidoarjo Regency. Nationally, the mapping method using ensemble Rock analysis will be applied in all districts in Indonesia so that each district can find out/map the status of the achievement of the Village SDGs which can be used as material for planning and evaluating village development.

Literature Review

Theoretical and Conceptual Background

About Village SDGs

SDGs are a global development agenda guided by five principles that harmonize the concepts of economic, social, environmental and government, namely people, planet, prosperity, peace, and partnership. (UN, 2015). Indonesia became one of the countries that agreed to this and it was stated in Presidential Regulation No. 59 of 2017 concerning the Implementation of Achieving Sustainable Development Goals. The SDGs have 17 goals measured by 169 indicators that are expected to be achieved by 2030. Based on the SDGs formulation, Indonesia has initiative to localized SDGs into the village level in order to solicit contributions and participation from many parties and to accelerate the achievement of targets in 2030 in Indonesia. The Village SDGs has 18 goals, it was built on the same five principles as the national SDGs, with the addition of one principle into six principles, namely Prosperous Villagers (People), Natural Balance (Planet), Village Prosperity (Prosperity), Peace, Partnership, Institutional and Village Wisdom. The implementation of the Village SDGs is aimed at directing development at the village level where to achieve these goals, the Village SDGs are prepared with 18 objectives (Iskandar, 2020). The Village SDGs are divided into 8 groups of the goals which are then referred to as fields. The 8 fields are poverty and hunger, health, education, women's friendliness, economy, environment, networked villages, and cultural responsiveness:

- i. In the Poverty and Hunger Sector there are 2 SDGs goals, namely Goal 1; Villages Without Poverty, and Goal 2; The Village Without Hunger.

- ii. In the Health sector, there are 3 goals, namely Goal 3; Healthy and Prosperous Villages, Goal 5; Decent Villages with Clean Water and Sanitation, and Goal 11; Safe and Comfortable Village Settlement Areas.
- iii. In the field of education, there is 1 SDGs goal, namely Quality Village Education.
- iv. In the Women-Friendly Sector, there is 1 goal of SDGs, namely the Involvement of Village Women.
- v. In the economic sector, there are 4 SDGs Goals, namely Goal 8; Equitable Village Economic Growth, Goal 9; Infrastructure and Village Innovation According to Needs, Goal 10; Villages Without Gaps, and Goal 12; Consumption and Production of Environmentally Aware Villages.
- vi. In the environmental sector, there are 4 SDGs Goals, namely Goal 7; Villages with Clean and Renewable Energy, Goal 13; Villages Responsive to Climate Change, Goal 14; Villages Caring for the Marine Environment, and Goal 15; Villages Caring for the Land Environment.
- vii. The Networking Sector consists of 1 Goal, namely Goal 17; Partnerships for Village Development.
- viii. The Cultural Response Sector consists of 2 objectives, namely Goal 16; Peaceful Villages with Justice and Goal 18; Dynamic Village Institutions

Measurement of SDGs Achievement

Measuring of SDGs is a topic that is very often discussed by researchers, especially regarding how to measure the performance of the SDGs. From a technical point of view, there is a plurality of frameworks, possible interpretations and selection of variables that includes conventions about the “correct” or “objective” method for measuring SDG performance. Miola and Schiltz (2019) discuss the differences in the three main methods used to measure SDGs in the European Union (EU). The first method is the Simple Mean method, which is a direct approach to building a measure of SDG performance by calculating the average of all indicators at the objective level. Thus, indicators are given the same weight and all weights apply to all countries – i.e. countries are not allowed to assign based on indicator weights more than others. Before calculating the average, the indicator needs to be re-scaled to calculate a different unit of measure. The second method is the Distance Measure method provided by EOCED's distance measure (OECD, 2017), the approach is a standardized calculation of the distance to the target specified for each indicator. And lastly is the measurement of progress where this method is a method based on the monitoring report by Eurostat on progress towards the SDGs in the EU context (2018). a value of 2030 is constructed for each country under the 'business-as-usual' (BAU) assumption. The BAU growth rate is calculated by extracting the first and last observations for each indicator and by linearly interpolating the final values in 2030. The resulting values for 2030 are combined into goals with the same feature scaling and weighting. The results of this paper indicate that whether or not the method used is based on the conditions of the country because each country measures different things so that the indicators used are also different. Different sets of indicators imply different priorities. Therefore, we argue that supporting indicators of country context should be used as they provide a more consistent source of data for country development. This research also finds important gaps in each of the analyzed methods. Neither of these methods considers interrelationships. Although this topic is very complex the holistic nature of the 2030 Agenda makes it a key element of any SDGs implementation policy, thereby placing more emphasis on contextualizing the choice of SDGs performance methods. (Miola & Schiltz, 2019).

Mantlana 2019 mapping the SDGs to understand the interrelationships between various SDGs targets. The results of this study indicate that all 17 goals of the SDGs interact with each other because they are basically interdependent. These results provide evidence that in order for the 2030 SDGs to be achieved, the implementation of the SDGs must be carried out as a mutually supportive unit to optimize benefits. In this research on the measurement of Village SDGs in Indonesia, the grouping of objectives is based on (Iskandar, 2020) which is divided into 8 groups as described in sub-chapter A, regarding Village SDGs.

Measurement of Achievement of Village SDG Goals in Indonesia

The implementation of the village SDGs measurement includes 18 goals which are described in targets and indicators. The Village SDGs consist of 222 indicators with 210 indicators of which are the Global and National SDGs indicators that can be applied in villages, and the other 12 indicators are the 18th Village SDGs indicators according to the results of village development research. The process of measuring the achievement of the Village SDGs based on data for the Village SDGs. According to the Regulation of the village minister (Examination with a Specific Purpose) No. 21/2020, Village SDGs data belongs to the village, so that the Village SDGs data collection is carried out using the participatory census method. The participatory dimension increases the validity of the Village SDGs data. The census dimension means taking data from all villages and neighborhood association areas, as well as collecting data from all families and villagers. (Kemendes PDDT, 2020). In detail, the instrument used for data collection are includes: a village questionnaire instrument filling by the village officials, neighborhood association questionnaire instrument, filling by the head of neighborhood association, family questionnaire instrument and citizen questionnaire instrument. Data processing and analysis is carried out electronically by the Village Information System developed by the Ministry of Villages, Development of Underdeveloped, and Transmigration. Further processing in the Village Information System produces recommendations for improving village development and community empowerment in order to achieve the goals in the Village SDGs. (Kemendesa, 2021).

Cluster Analysis

Cluster analysis or object grouping is a statistical technique for grouping objects that have similarities in one group or cluster and which have differences with other groups. (Johnson & Wichern, 2007). Group analysis is distinguished into group analysis of numerical and categorical data as follows.

Numerical Data Cluster Analysis

Numerical data cluster analysis used in this study is the agglomerative hierarchical method. The distance function that is often used is the Euclidean distance. Euclidean distance is the distance between objects to measure the similarity with equation (1).

$$d_{ij} = \sqrt{\sum_{h=1}^p (X_{ih} - X_{jh})^2}, i \neq j$$

Single linkage method single linkage is a hierarchical cluster method that is formed based on the distance or similarity between pairs of objects. Determination procedure:

1) single linkage cluster that is through the shortest distance; 2) complete linkage, ie all objects in a group are linked to each other based on the most distant similarity; 3) Average Linkage hierarchical cluster analysis method which treats the distance between two clusters as the average distance between all individual pairs. So the distance between (UV) and every other W cluster (Johnson & Wichern, 2007)

Categorical Data Cluster Analysis

The ROCK (Robust Clustering with Links) method uses the link concept to measure the similarity or closeness between observations.

The number of links between observations depends on the threshold value (θ) that has been determined. Value of θ is a parameter to state the existence of a link between observations. If the distance between a pair of observations is greater than the value of θ , then the observations have a (Guha, Rastogi, & Shim, 2000).

The ROCK method has 3 steps, namely:

- Calculating similarity with following equation:

$$sim(X_i, X_j) = \frac{|X_i \cap X_j|}{|X_i \cup X_j|}, i \neq j$$

- Determine neighbors, observations are declared as neighbors if the value;

- Calculate the link between the object of observation. The link size is influenced by the threshold value (θ) which is used to control how close the relationship between objects is where. Merging groups using the ROCK algorithm is based on the goodness measure between groups, namely the number of links divided by the possible links formed based on the size of the group (Tyagi & Sharma, 2012).

Mixed Data Cluster Analysis

The results of grouping numerical and categorical data are combined using the ROCK ensemble to obtain the final group (He, Xu, & Deng, 2005). The ensemble grouping scheme for mixed data can use the CEBMDC (Cluster Ensemble Based Mixed Data Clustering) algorithm with the following steps.

- i. Divide the data into two subdata, namely purely numeric and purely categorical.
- ii. Grouping objects that have numeric variables with numeric data grouping algorithms, and grouping objects that have categorical variables using categorical data grouping algorithms.
- iii. Combining the grouping results of numeric and categorical variables, which is called the ensemble process.
- iv. Perform ensemble grouping using categorical data grouping algorithms to get the final cluster.

Performance of Grouping Results

Measuring the performance of clustering results is a step to determine the validity of a cluster. A good cluster will have high homogeneity between members in the group and high heterogeneity between members outside the group (Hair, Black, Babin, & Anderson, 2010). grouping using Pseduo-F value and ICD Rate. The highest pseudo F indicates that the group provides optimal results. The performance of clustering results for variables with a numerical data scale is shown in the following equation:

$$Pseudo F = \frac{\left(\frac{R^2}{k-1} \right)}{\left(\frac{1-R^2}{n-k} \right)}$$

$$ICD\ rate = 1 - \frac{SSB}{SST} = 1 - R^2$$

where,

$$R^2 = \frac{(SST - SSW)}{SST} = \frac{SSB}{SST}$$

$$SST = \sum_{i=1}^n \sum_{h=1}^p \sum_{k=1}^g (x_{ihk} - \bar{x}_h)^2$$

Sum Square Total =

$$SSW = \sum_{i=1}^n \sum_{h=1}^p \sum_{k=1}^g (x_{ihk} - \bar{x}_{hk})^2$$

Sum Square Within Group =

The smaller Icdrate value indicates that the cluster is getting better because members in one cluster have low differences and have small variations. (Mingoti & Lima, 2006). The performance of clustering results for variables with categorical data scale is shown in the following equation:

$$SST = \frac{n}{2} - \frac{1}{2n} \sum_{k=1}^K n_k^2$$

Sum Square Total =

Sum Square Within Group (SSW)

$$SSW = \sum_{c=1}^C \left(\frac{n_c}{2} - \frac{1}{2n} \sum_{k=1}^K n_{kc}^2 \right) = \frac{n}{2} - \frac{1}{2} \sum_{c=1}^C \frac{1}{n_c} \sum_{k=1}^K n_{kc}^2$$

Sum Square Between Group (SSB)

$$SSB = \frac{1}{2} \left(\sum_{c=1}^C \frac{1}{n_c} \sum_{k=1}^K n_{kc}^2 \right) - \frac{1}{2n} \sum_{k=1}^K n_k^2$$

$$MST = \frac{SST}{(n-1)}$$

Mean of Squares Total =

$$MSW = \frac{SSW}{(n-C)}$$

Mean of Squares Within =

$$MSB = \frac{SSB}{C-1}$$

Mean of Squares Between =

The standard deviation within groups (SW) and standard deviation between groups (SB) for categorical data can be the roots of MSW and MSB. The performance of a grouping with categorical data is based on a comparison of the ratio between the standard deviations within the group (SW) and the standard deviations between groups (SB). If the comparison ratio is getting smaller then the performance of grouping categorical data is getting better because of the maximum homogeneity in the group and maximum heterogeneity in the group.

Research & Methodology

Data Source

The data used in this study is secondary data from the Department of Community and Village Empowerment of Sidoarjo Regency. The data is data from the Sidoarjo Regency Village Development Index (IDM) survey in 2020 which was then adjusted to the indicators from the Village SDGs. The sample unit used is a village, with a total of 322 villages in Sidoarjo Regency.

Concepts and Variables

Village Clustering based on the Village SDGs Indicators of the Economic Aspect.

The Village SDGs in the economic sector consist of 4 indicators, namely indicators of equitable economic growth, indicators of infrastructure and innovation as needed, indicators of villages without inequality, indicators of consumption and environmentally conscious production.

Grouping of Village SDGs in the Environmental Care Sector

The Village SDGs in the field of environmental care consist of 4 indicators, namely clean and renewable energy village indicators, climate change responsive village indicators, village environmental care indicators, and partnerships for village environmental development.

Grouping of Village SDGs in Health and Education Sector

The grouping of villages based on the village SDGs indicators in the health sector includes grouping based on indicators of healthy and prosperous villages, decent sanitation villages, safe and comfortable village residential areas. In this subsection, villages will also be grouped based on Education indicators.

Grouping of Village SDGs in the Sector of Community Welfare Equity

In the field of equitable distribution of community welfare, it consists of the field of poverty and hunger, the field of women's friendliness, and the field of cultural responsiveness.

The variables used are 18 Village SDGs except Goal number 14, namely Village Care for the Marine Environment, the variables used consist of 76 variables that refer to the objectives of the Village SDGs, shown in Table 1.

Table 1: Research Variables

Goal	Variables	Scale
1	Number of Poor Families, number of SJSN recipients,	Numeric
2	Percentage of population with malnutrition, stunting	Numeric
3	Percent of BPJS, Number of MMR, Number of IMR, Percent of Immunization, Posyandu 2 months, Number of Health Workers, Number of Health Facilities, BPD KLB	Numeric
4	APM preschool, Number of Preschool POS, Number of Junior High School Illiterate, Write Read and Count	Categorical
5	Head of PR , Secretary of PR	Numerik
6	Defecate in private latrines, defecate together in latrines, defecate in public latrines	Categorical
7	Access to electricity, solar energy, wind, fuel use, base	Numeric
8	The main source of village income, superior products, types of superior products, KUR,	Categorical
	Active Cooperatives, Number of Economic Infrastructure Courses, Financial Institutions, Number of Village-Owned Enterprises	Categorical
9	Road Surface , Road Quality	Numeric
	Total City Minimum Wage, Total Minimum Wage	Numeric
10	Poor Family, Non-Permanent House	Categorical
11	Disaster Early Warning, R. public Environmental security post	Categorical
12	Waste Management, waste disposal into rivers, Waste Management	Categorical
13	Irrigation, disaster events	Categorical
14	<i>Not Analyzed</i>	
15	water pollution, soil pollution, air pollution, topography	Categorical
16	Crime, Trafficking, Suicide	Categorical
17	BTS tower, post office, expedition service, internet cafe, signal, internet at village office, village website, other information facilities	Categorical
18	Secretariat, Technical Executor, Territory Implementation, Institution, Cohesiveness, birth, death, marriage, village-owned enterprises	Categorical

Analysis Steps

Based on the data sources and research variables that have been described previously, the following analysis steps were carried out in this study consisting of 5 stages, namely:

- Data preparation (starting with sorting numerical and categorical data for each aspect studied)
- Village grouping for numerical data is done using agglomerative hierarchical method with the optimum cluster results on the average linkage method for the highest Pseudo-F value of 9.570948 and the smallest ICD Rate.

- iii. Analyze categorical data using the ROCK method.
- iv. Combine the results of the hierarchical and ROC methods, namely the ROCK ensemble method.
- v. Summarize the cluster results from each Village SDGs goal into 1 final cluster.

Problem Limitation

In this study, data sourced from the IDM (Development Village Index) of Sidoarjo Regency was used in 2020. The issue that is still an issue is related to the completeness and accuracy of the SDGs Village data.

Analysis and Findings

Villages Clustering based on Indicators in Economic Sector

Using agglomerative hierarchical method for numerical variables, the optimum cluster results is 4 clusters with the highest Pseudo-F value of 9.570948 and the smallest ICD Rate. The four clusters consist of: 318 villages in the first cluster, 1 village in the second cluster, 2 villages in the third cluster and 1 village in the fourth cluster.

While for categorical data, using the ROCK method at threshold of 0.05, 0.10, 0.15, 0.20, 0.25, and 0.30 produces the optimum cluster based on the smallest SW and SB ratio. For instant at the threshold of 0.30 with the number of clusters formed is 2 clusters, consisting of 111 villages in cluster 1 and 211 villages in cluster 2.

The best cluster results from the 2 methods were regrouped using the ROCK ensemble method. The results showed that the ratio between the smallest SW and SB values was at the 0.30 threshold with the number of clusters formed 2 clusters consisting of 82 villages in cluster 1 and 240 villages in cluster 2 shown in Figure 1.

Figure 1: Village mapping based on Economic aspect

In cluster 1 village, the average number of active cooperatives is 2 active cooperative facilities, the average of availability of financial institutions and businesses higher than cluster 2. Cluster 1 has the highest average number of poor families compared to cluster 2 with number of non-permanent houses of 11 houses in average. In cluster 1, most villages dispose of liquid waste in drainage and landfills. Garbage is a trash can which is then transported. In addition, cluster 1 tends to have BTS towers, post offices, expedition services, internet cafes, strong category telephone signals, village web presence and other superior information facilities.

Economic facilities are mostly located in cluster 2 with a maximum number of 420 facilities in one village Micro and small industries have the highest maximum value in cluster 2, which is 679 industries in one village and medium industries with a maximum number of 685 industries in another village. The village's main income in cluster 2 is mostly from agriculture. In addition, the existence of post offices and expedition services tends to be quite low. So that cluster 1 is included in the category of high economic sector while cluster 2 is classified in the category of low economic sector.

In the same way for the other 3 aspects, the following results are obtained:

Village Clustering based on Indicator of Environmental Indicators

Analysis of mixed data grouping using the ensemble ROCK method found that the smallest ratio value was 0.0071 at the threshold values of 0.25 and 0.30 and the optimum number of clusters was 2 clusters. Cluster 1 has 241 villages and cluster 2 has 81 villages, the village mapping is presented in Figure 2.

Figure 2: Clustering of Villages Based on Environmental

In the village indicators care for the environment, it is known that the highest average percentage of electricity user families is in cluster 1, which is 94.946% with an average percentage of electricity user families reaching 100% in cluster 1 of 56.62%, which is superior to the percentage in cluster 2. Cluster 3 consists of 158 villages, excels in the complete basic immunization variable and is a village group with moderate health levels.

The distribution of villages included in each cluster is presented in Figure 3.

Figure 3: Health Sector Village Grouping

Village Clustering Based on Education Indicators

The results of grouping villages in Sidoarjo Regency based on the SDGs indicators in the education sector using the ensemble ROCK method obtained 3 optimum clusters, namely: cluster 1 there are 60 villages, cluster 2 there are 162 villages, and cluster 3 there are 100 villages.

The distribution of villages included in each cluster is presented in Figure 4 which provides information that the yellow color is the villages that are included in cluster 1, the green color is the villages that are included in cluster 2, and the brown color is the villages that are included in the cluster 2. cluster 3. Nearby villages tend to have the same characteristics, namely being in the same cluster membership.

Cluster 1: has the largest population, most of the population graduated from elementary school/equivalent at most the kindergarten NER value is higher than the other clusters, which is around 73.68%. The number of preschool posts is higher than clusters 2 and 3, where the number of preschool POS in cluster 1 ranges from 0 to 20 with an average of about 6 units per village. The existence of illiteracy eradication activities and the existence of A/B/C package learning centers are also superior in cluster 1. So that cluster 1 is a group with higher education concerns.

Figure 4: Village Grouping in Education Sector

Cluster 2: excels in terms of the availability of community reading gardens (94% of 162 villages in this cluster already have a reading garden). Cluster 2 as a village group with moderate education concerns, Cluster 3: does not have an advantage, where there are 80% of villages that do not yet have a community reading park. Cluster 3 is a village group with low education awareness.

Village Clustering based on Equitable Community Welfare Indicators

In the field of equitable of community welfare, it consists of the poverty and hunger, women's friendliness, and cultural responsiveness.

Village Clustering Based on Poverty and Hunger Indicator

The clustering of villages in poverty and hunger indicators using the agglomerative hierarchical cluster method showed that the complete linkage has the largest Pseudo F value and the smallest ICD Rate with a number of clusters of 5 so that the optimum cluster formed is 5 groups. After calculating the characteristics of each cluster with the mapping according to Figure 5.

Figure 5: Village Grouping in Poverty and Hunger Sector

Cluster 1 can be identified as a village group with a high BPJS health participation rate,

Cluster 2 is poor on the percentage of poor families, is a village group with a high level of poor families

Cluster 3 is poor in the percentage of health insurance participants and the percentage of stunting people is quite high

Cluster 4 is group of village with moderate levels of malnutrition.

Cluster 5: consists of 1 village with the highest percentage of malnutrition, which is 38.46%.

Village Clustering Based on Gender-equality Indicator

Gender-equality measured by the involvement of women in the presence of village management. The variables used have categorical tipe so that its clustered using the ROCK method. Based on the ROCK analysis, it was found that the lowest ratio value is 0.0175 at the threshold of 0.05 and 0.10. The clusters formed at the two thresholds have the same members so with an optimum cluster of 2. It was found 246 villages in cluster 1, as a village group that does not have women's participation in village government positions.

Cluster 2 consist of 76 villages, is a village group that still has women's participation in village government positions. Mapping of villages according to the clusters formed is shown in Figure 6.

Figure 6: Clustering of Villages Based on Gender -Equality

Village Clustering Based on Cultural Responsiveness

The clustering of villages in the cultural responsiveness aspect have been analysed by the ROCK Ensemble method because it has numeric and categorical type variables. It was found that the lowest ratio value was found at a threshold of 0.15 to 0.30 which was 0.0000 but at that threshold it could not form a cluster so that the threshold value used was 0.55 with a ratio of $9,66 \times 10^{-10}$ which formed 2 groups.

Figure 7: Cultural Response Village Grouping

Figure 7 shows that cluster 1 was identified as a village group with high cultural preservation and cluster 2 was a village group with low cultural preservation.

Cluster Analysis of Village based on the Achievement of Village SDGs

Figure 8 shows the results of cluster analysis based on 18 SDGs goals except for the 14th goal, namely marine ecosystems. It can be seen that the best achievement is goal number 6, Village Adequate Sanitation, about 319 villages are classified in cluster 1 (the best cluster). The average percentage of families who have access to a latrine is quite large, namely 90.93%. The second best achievement is the 2nd goal, Village without Hunger. About 317 villages in Sidoarjo Regency have a very low average percentage of malnourished children under five, only 1.745% and an average percentage of stunting under five is 2.414%.

On the other hand, goal number 18 (Dynamic Village Institutions and Adaptive Village Culture) also have good achievement. About 92% villages in Sidoarjo Regency have also been included in cluster 1. The achievement of goal number 10 (Village Inequality), Goal 9 (Infrastructure and Village Innovation as Needed), and Goal 3 (Healthy and Prosperous Village) Sidoarjo Regency also needs to improve several goals where there are still very few villages that are in Cluster 1 or the best Cluster, namely Goal 13 (Climate Change Responding Village) and Goal 7 (Clean and Renewable Energy Village).

Figure 8: Resume Grouping Based on 18 SDGs

Conclusion

Clustering using Ensemble ROCK method is suitable for village SDGs achievement, it can be accommodated the two type of data od DSGs measurement numerical and categorical data. The overall result af Ensamle ROCK analysis can express the performance of Village in term of SDGs achievement.

It can be concluded that among 322 villages in Sidoarjo Regency, 82 villages have the high economic field, 241 villages are superior in environmental field, 158 villages having medium health aspect, 60 villages belonging to the village in the field of higher education and 162 villages belonging to the village in the medium education sector, 235 villages identified as a village group with a high health assurance participation rate. On the other hand, 240 villages having low economic field, 81 villages having low environmental concern, 100 villages are low in education, 163 villages with low health sector.

In the case of Community Welfare Equity, 76 villages having a high level of poor families, 4 villages with a high stunting rate, 6 villages with moderate levels of malnutrition, and 1 village with high levels of malnutrition.

Intermf of women-participation, 246 villages classified as a village group with no female participation in village government positions and 76 villages classified as a village group with no female participation in village government positions. In the field of cultural responsiveness, almost all villages to be identified having high cultural preservation.

Based on 18 goals of Village SDG, more than 60% village can achieve 8 goals, namely village without hunger, healthy and prosperous village, village adequate sanitation, village infrastructure and innovation as needed, village without gaps, safe and comfortable village residential area, consumption and production of environmentally aware villages, dynamic village institutions and adaptive village culture. It also can be concluded that Sidoarjo Regency still has to find appropriate strategy to achieve goal 1 (Village Without Poverty), goal 4 (Quality Village Education), goal 5 (Involvement of Village Women), goal 7 (Clean and Renewable Energy Village), goal 13 (Climate Change Responsive Villages).

Sidoarjo Regency is expected to pay more attention to villages that are still lagging behind based on the Village SDGs target. The majority of villages in Sidoarjo Regency are still lagging behind in the fields of economy, health, education, and community welfare.

Clustering of the achievement of Village SDGs is very useful tools to understand the status of the village achievent. Mapping of SDGs achievement in village level is very important information-built strategy and program for regional (Regency) development.

Author Contributions: Conceptualization, ATR, NZS, ARN, HJS, LFR.; Methodology, ATR, NZS, ARN, HJS, LFR.; Data Collection, ATR, NZS, ARN, HJS, LFR.; Formal Analysis, ATR, NZS, ARN, HJS, LFR.; Writing—Original Draft Preparation, ATR, NZS, ARN, HJS, LFR.; Writing—Review And Editing, ATR, NZS, ARN, HJS, LFR. All authors have read and agreed to the published the final version of the manuscript.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to that the research does not deal with vulnerable groups or sensitive issues.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Alvionita. (2017). Metode Ensembel ROCK dan SWFM untuk Pengelompokan Data Campuran Numerik dan Kategorik pada Kasus Akses Jeruk. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Bappenas. (2017). Ringkasan Metadata Indikator TPB/SDGs Indonesia. Jakarta: Bappenas.
- DINPMD. (2021). Dinas Pemberdayaan Masyarakat dan Desa Kabupaten Sidoarjo. Retrieved Juni 15, 2021, from <http://sid.sidoarjokab.go.id/>
- Guha, S., Rastogi, R., & Shim, K. (2000). ROCK: A Robust Clustering Algorithm for Categorical Attributes. Proceeding of the 15th International Conference on Data Engineering.
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). Multivariate Data Analysis. United State of America: Pretice Hall International, Inc.
- He, Z., Xu, X. i., & Deng, S. (2005). Clustering Mixed Numeric and Categorical Data: A Cluster Ensemble Approach. Departemen of Computer Science and Engineering, Harbin Institute of Technology.
- Iskandar, A. (2020). SDGs Desa Percepatan Pencapaian Tujuan Pembangunan Nasional Berkelanjutan (1st ed). Jakarta: Pustaka Obor.
- Iskandar, A. H. (2020). SDGs Desa Percepatan Pencapaian Tujuan Pembangunan Nasional Berkelanjutan (1st ed.). Jakarta: Pustaka Obor.
- Johnson, R., & Wichern, D. (2007). Applied Multivariate Statistical Analysis 6th-Edition. New Jersey: Pearson Prentice Hall.
- KDPDPTT. (2021, Juli). Tentang Indeks Desa Membangun. Retrieved from Direktorat Jenderal Pembangunan dan Pemberdayaan Masyarakat Desa: <https://idm.kemendes.go.id/view/detil/1/tentang-idm>
- Kemendes PDTT. (2020). Permendes PDTT No 21 Tahun 2020 Tentang Pembangunan dan Pemberdayaan Masyarakat Desa. Jakarta.
- Kemendesa. (2021, 02 22). Pendataan SDGs Desa 2021. Retrieved from <https://sdgsdesa.kemendes.go.id/:https://sdgsdesa.kemendes.go.id/pemutakhiran-data-sdgs-desa/>
- Mingoti, S., & Lima, J. (2006). Comparing SOM neural network with Fuzzy c-means. European Journal of Operational Research, 1748.
- Miola, A., & Schiltz, F. (2019). Measuring sustainable development goals performance: How to monitor policy action in the 2030 Agenda implementation? Ecological Economics.
- OECD. (2017). Measuring Distance to the SDG Targets: An Assessment of Where OECD Countries Stand. Paris.
- PMD. (2021, Juni). Penghargaan Utama SDGs Dari Kemendes PDTT Untuk Desa Yang Selesaikan Pendataan Tepat Waktu. Retrieved from Dinas Pemberdayaan Masyarakat Dan Desa Kabupaten Sidoarjo: <http://sid.sidoarjokab.go.id/2021/06/09/penghargaan-utama-sdgs-dari-kemendes-pdtt-untuk-desa-yang-selesaikan-pendataan-tepat-waktu/>
- Ramdhany, R. (2018). Pengelompokan Desa Di Kabupaten Bondowoso Berdasarkan Data Campuran Numerik Dan Kategorik Menggunakan. Surabaya: Statistika, Institut Teknologi Sepuluh Nopember.
- Tyagi, A., & Sharma, S. (2012). Implementation of ROCK Clustering Algorithm for the Optimaztion of Query Searching Time. International Journal on Computer Science and Engineering, 05.
- UN. (2015). Transforming Our World: The 2030 Agenda for Sustainable Development. New York: United Nations.

Publisher's Note: SSBFNET stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2022 by the authors. Licensee SSBFNET, Istanbul, Turkey. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

International Journal of Research in Business and Social Science (2147-4478) by SSBFNET is licensed under a Creative Commons Attribution 4.0 International License.