

Efforts to Alleviate Underdeveloped Areas by Clustering Regional Characteristics in Indonesia

Suyanto^{1*}, Wiwik Budiarti², Rahmawati Erma Standsyah³, Dendy Syahru Ramadhan⁴

Universitas Dr. Soetomo, Semolowaru 84 Surabaya

Corresponding Author: Suyanto suyanto_fe@unitomo.ac.id

ARTICLE INFO

Keywords: Alleviate Underdeveloped Areas, Clustering, Economic

Received : 18, October

Revised : 22, November

Accepted: 24, December

©2023 Suyanto, Budiarti, Standsyah, Ramadhan: This is an open-access article distributed under the terms of the [Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



ABSTRACT

This study aims to cluster the underdeveloped regions in Indonesia according to the criteria of the underdeveloped indicator to mitigate the underdeveloped regions in Indonesia. This research was conducted to help the various efforts made by the government to deal with the underdeveloped regions, by grouping the underdeveloped regions, it is hoped that the government can focus on increasing the dominant criteria in the regions according to the cluster. The grouping method used is K-means with the results of 62 underdeveloped districts in Indonesia divided into 3 clusters. The first cluster includes 23 districts grouped based on human resource criteria, the second cluster consists of 28 districts based on infrastructure/facilities criteria, and the third cluster consists of 11 districts based on economics criteria.

INTRODUCTION

In 2045, Indonesia will be 100 years after its independence in 1945. This is what drives Indonesia to improve itself to overcome economic, social, and other problems in this country. One of the problems that the government is currently in charge of is the problem of underdeveloped areas. In 2020, the government issued Presidential Decree No. 63 of 2020 regarding the determination of disadvantaged areas. This regulation states that there are 62 regencies spread over 10 provinces in Indonesia as underdeveloped regions. The determination of these regions is based on several predetermined criteria (Presiden of the Republic Indonesia, 2020).

The problems regarding the underdeveloped areas have existed in Indonesia for a long time but have not been completely resolved. The government has made several efforts to reduce the existing disadvantaged areas, one of the efforts made by the president is the Periphery Development Program which is included in the development agenda in the National Medium Term Development Plan 2020-2024 (Presiden of Republic Indonesia, 2020). This program achieves sustainable development in rural areas by involving all village potentialities, the objective is to increase the dissemination of development. The Ministry of Villages, through the Directorate General of Underdeveloped Areas Development, has launched the Development Management Information System application, with the objective that within the next five years, 10,000 underdeveloped villages can become developing villages and 5,000 developing villages into independent villages.

Several researchers have also studied how to deal with underdeveloped areas such as the Estimation of Poverty Model Effect Size in Indonesia Using Meta-Analytic Structural Equation Modeling (MASEM) by (Standsyah et al., 2023) which discusses the relationship between variables in underdevelopment criteria, Assessment and key factors of urban liveability in underdeveloped regions: A case study of the Loess Plateau, China by (Xiao et al., 2022) which analyzes important factors in the existence of underdeveloped regions, and Unleashing the potential of mobile broadband: Evidence from Indonesia's underdeveloped regions on its role in reducing income inequality by (Ariansyah et al., 2023) which examines the relationship between the development of information technology and the society and economy in underdeveloped areas.

Some researchers have also studied underdeveloped areas such as (Otok et al., 2021) with A meta confirmatory factor analysis of the underdeveloped areas in the Java Island, Green infrastructures for urban sustainability: Issues, implications, and solutions for underdeveloped areas by (Cheshmehzangi et al., 2021), Determinants of Innovation Ecosystem in Underdeveloped Areas – Take Nanning High-Tech Zone in Western China as an Example by (Huang et al., 2020), (Wu et al., 2023) with Uncertainty analysis of envelope retrofits for existing residential buildings in underdeveloped areas: A case study of Daokou, China, (Yu et al., 2013) with Factors Impact The Doot-to-Balloon Time in Stemi Patients in an Underdeveloped Area of China: a Single-Centered Analysis, Regional governments and opportunity entrepreneurship in underdeveloped institutional environments: An entrepreneurial ecosystem perspective by (Wei,

2022), Prospects for the Development of Underdeveloped Territories on the Basis of Energy and Transport Infrastructure by (Prokhorov et al., 2023).

The various efforts that have been made before have not yet given optimal results in solving the problems of underdeveloped regions, that's why this is a research subject that must be developed. To overcome the problem of underdeveloped areas, it is necessary to know the root causes of these areas, underdeveloped regions themselves are districts/cities whose territories and people are less developed compared to other regions at a national scale. The area can be said to be underdeveloped because it does not meet several criteria such as (1) Economic criteria, (2) Human resources criteria, (3) Regional financial capacity criteria, (4) Infrastructure/Facilities Criteria, (5) Accessibility Criteria, (6) Criteria for regional characteristics (Regulation of the Minister of Village, Development of Underdeveloped Areas, and Transmigration of The Republic Indonesia No. 6, 2016). Several researchers in Indonesia have also studied the problems of underdeveloped regions such as Modeling Underdeveloped Regions in Indonesia Using Discriminant Analysis (Purwandari et al., 2017), Analysis of Determining Underdeveloped Regions for 2020-2024 and the National Action Plan for Accelerating the Development of Underdeveloped Regions for 2020 (Nabila et al., 2021).

In statistics, clustering is divided into two methods i.e. clustering with hierarchical approach i.e. clustering data by creating a hierarchy in the form of a dendrogram where similar data will be placed in adjacent hierarchies and those not in distant hierarchies, and clustering with a partition approach, which will cluster data by sorting the analyzed data into multiple existing clusters. This ignores the hierarchy of data, unlike the hierarchical approach. The K-means clustering method is included in the method with the partitioning approach, the K-means method has the advantages of being relatively simple and easy to implement and can be scaled to large data sets. Some examples of researchers using other partitioning approaches are Clustering of Underdeveloped Area Infrastructure with an Unsupervised Learning Approach: A Case Study in the Island of Java, Indonesia (Otok et al., 2022), and An effective partitioning clustering algorithm based on new clustering validity index (Zhu et al., 2018).

The K-Means clustering method is a non-hierarchical method that will group data into a predetermined number of clusters, with each data in the cluster having the same characteristics (Hu et al., 2023; Ikotun et al., 2023). The variables that will be used are underdeveloped criteria comprising economic criteria, HR criteria, as well as facilities/infrastructure criteria according to their respective indicators with 62 regency/city units that are classified as underdeveloped areas.

This research will focus on the grouping of underdeveloped areas in Indonesia which will be divided into 3 groups based on the underdeveloped criteria. In order to be able to group the districts based on underdeveloped indicators, and to discover areas with criteria that must be prioritized to be treated first.

LITERATURE REVIEW

The data used in this study were obtained from the Indonesian Central Bureau of Statistics. The data used is in the form of underdevelopment indicators for districts/cities in Indonesia, which consist of 62 underdeveloped districts/cities in Indonesia in 2021 (Presidential Regulation of Republic Indonesia No. 63, 2020) (D. of D. R. A. T. of the R. of I. Ministry of Villages, 2016).

Underdeveloped Criteria and Indicators to identify a district/city experiencing underdevelopment can be measured using predetermined standards. Indicators of underdeveloped areas include:

- a. Economic Criteria. It consists of 2 indicators: (1) the amount of poverty and (2) the expenditure per capita of the population.
- b. Human Resource Criteria. Consist of 3 indicators: (1) Life Expectancy, (2) Average Years of Schooling, and (3) Literacy Rate.
- c. Infrastructure/Facilities Criteria. It consists of 3 indicators: (1) Average distance to the district capital, (2) Adequate water users, and (3) Proper sanitation users.

METHODOLOGY

K-Means Clustering Method

K-means is the most popular clustering method for use in various fields because the method is simple and easy to implement. K-means is a partitioning clustering method that separates data into different groups of clusters. The purpose of this data clustering is to minimize the objective function in the clustering process, which typically seeks to minimize within-group variance and maximize between-group variance (Soemartini et al., 2017) (M. W. Talakua et al., 2017). This method is used because this method will analyze and classify certain data with different variables, depending on the characteristics that exist in each cluster.

The K-means method is used as an alternative to the cluster method for relatively large amounts of data. Indeed, this method has a higher speed compared to the hierarchical method in general. The K-means method can be used to explain the algorithm for determining an object in a particular cluster based on the closest mean (Sani, 2018).

RESULT

We divide the clustering into 3 groups according to the underdevelopment criteria, so that we get the results of clustering where the first cluster with the most prominent variable is the HR criteria, the second cluster is the infrastructure criteria, and the third cluster is the economic criteria.

The number of Members of Clusters 1, 2, and 3 out of a total of 62 Regencies/Cities which are classified as underdeveloped regions in Indonesia are listed in Table 1. The results of the cluster division of all underdeveloped regions in Indonesia with their territories can be seen in the following Picture 1.

Table 1 Number of Each Member in each Cluster

Cluster Group	Number of Member
1	23
2	28
3	11
Total	62

Source: Data Processed

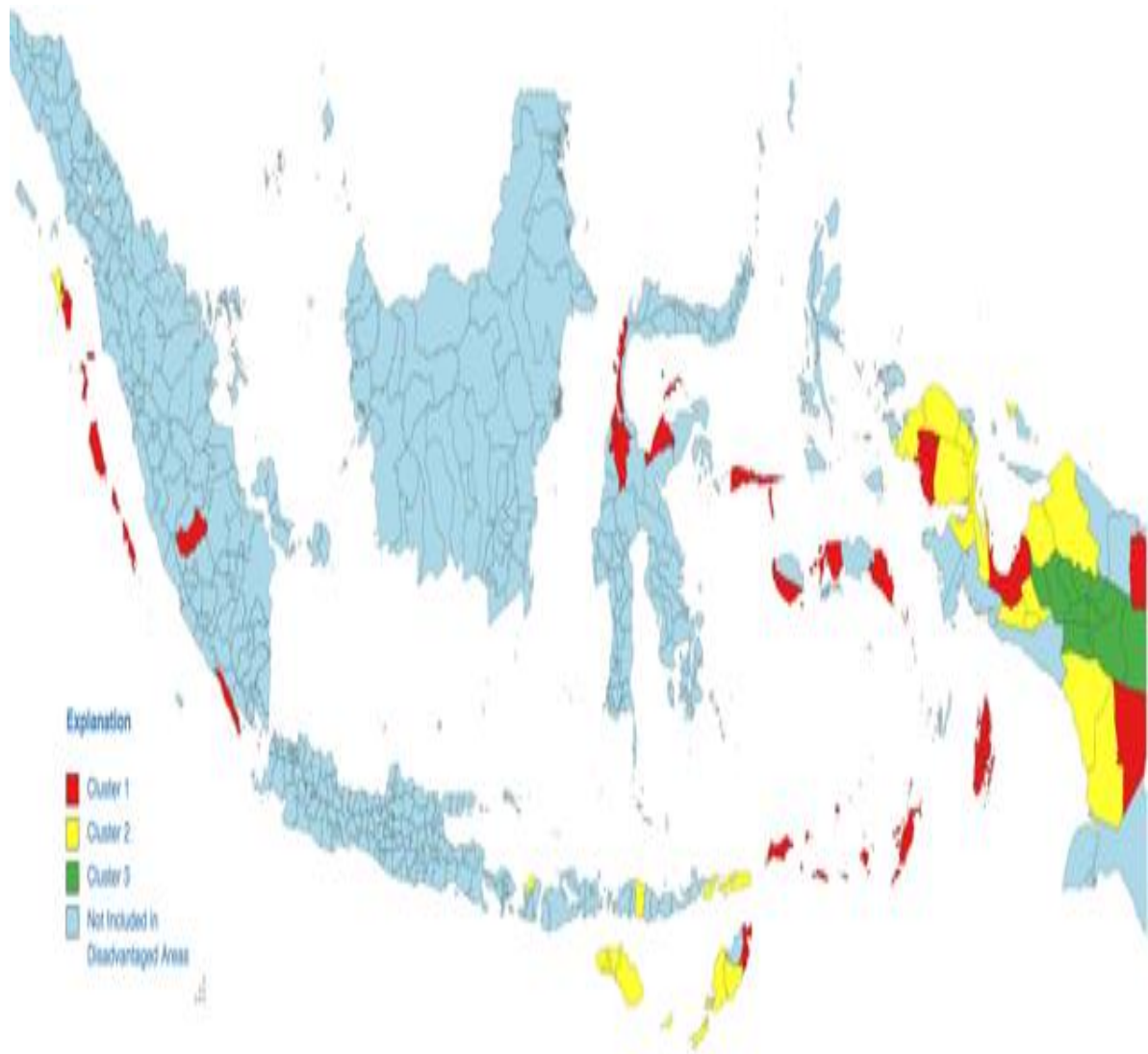


Figure 1. Cluster Division on Indonesia's Map Divided by Colours
Source: Data Processed

This research resulted in 3 regional groups in which each cluster was divided based on the criteria of Facilities and Infrastructure, HR Criteria, and Economic Criteria. Each group includes:

Cluster 1:

The first cluster shows that the variables Life expectancy, Average years of school, and Literacy rate have higher values than the other groups. While cluster 1 members include regency which is shown in Picture 1 with a red colored area, it means that the HR criteria of group 1 are the most important and it can be concluded that the regions mentioned above must optimize the HR development

The HR criteria consist of various indicators, such as life expectancy, the average length of schooling, and literacy rate. Life expectancy itself means the average number of years that a person can live from birth. The average length of schooling, as the name suggests, is the number of years that the population spends in formal education. While the literacy rate is the proportion of the 15 years and over of people who can read and write.

Cluster 2:

The second Cluster is based on the variables Distance to the Capital, Sanitation Access, and Adequate Water Access which have higher values than other clusters. While members of the second group include the regency colored yellow in Picture 1. This means that the criteria for Facilities/Infrastructure in Cluster 2 are the most prominent, which can be concluded if these areas can optimize the use of village funds according to the Criteria for Facilities/Infrastructure.

The Infrastructure/Facilities Criteria has several variables including distance to the district capital, adequate water users, and proper sanitation users. As the name suggests, the distance to the capital is the distance from a certain area to the provincial capital. Adequate water users are the percentage of the population of an area that uses proper water. And sanitation users are the percentage of the population using their own latrines (Andhika Arie Prasetya et al., 2021).

Cluster 3:

Unlike the other clusters, cluster three is grouped on the variable Percentage of Poverty and Per Capita Expenditures which have higher values than other clusters. While the members of the third group include the area shown in Picture 1 with green colored area, which means that the economic criteria in Cluster 3 are the most prominent, which can be concluded if these areas can optimize the use of village funds according to the Economic Criteria.

DISCUSSION

This economic criterion is divided into 2 indicators: the number of poverty, and the expenditure per capita of the population. The amount of poverty is the total of people that have an average per capita expenditure below the poverty line. Meanwhile, per capita expenditure is the monthly cost spent on average family consumption.

CONCLUSIONS AND RECOMMENDATIONS

From the clustering analysis study conducted using the K-means method above, it can be concluded that the underdeveloped areas included in the first group should prioritize improving their HR criteria, with the life expectancy, the average years of school, and literacy rate as its variable. Meanwhile, the second group should improve its infrastructure/facilities criteria with distance to the district capital, adequate water users, and sanitation users themselves as a variable, and the third group should advance its economic criteria with the amount of poverty, and expenditure per capita for its variable. With these results, the village can more easily determine the best use of village funds so that it can provide the effect the community needs most in that area.

The results of this study focus only on the grouping of underdeveloped areas, and further research can still be conducted using variables across all districts/cities in Indonesia. Using the same method with more data, it is necessary to multiply the cluster splits, so that the cluster split for each region can be even more accurate and have the appropriate characteristics for each cluster. Through this research, it is hoped that the government will be able to easily determine the needs of each region of Indonesia.

ADVANCED RESEARCH

The results of this study can be a reference for other research in the future and can be developed to make the results of renewable researchers in the form of scientific articles or others.

ACKNOWLEDGMENT

Gratitude and pride in the completion of the results of this research in the form of scientific articles, thank you infinitely to our university institutions for motivation, enthusiasm and these results can make a real contribution and can be useful university today or in the future.

REFERENCES

- Ariansyah, K., Barsei, A. N., Syahr, Z. H. A., Sipahutar, N. Y. P., Damanik, M. P., Perdananugraha, G. M., Dunan, A., Nupikso, D., Hidayat, D., & Mudjiyanto, B. (2023). Unleashing the potential of mobile broadband: Evidence from Indonesia's underdeveloped regions on its role in reducing income inequality. *Telematics and Informatics*, 82, 102012.
- Cheshmehzangi, A., Butters, C., Xie, L., & Dawodu, A. (2021). Green infrastructures for urban sustainability: Issues, implications, and solutions for underdeveloped areas. *Urban Forestry & Urban Greening*, 59, 127028.
- Huang, F., Cao, Z., Guo, J., Jiang, S.-H., Li, S., & Guo, Z. (2020). Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. *Catena*, 191, 104580.
- Nabila, S. P., Ulinnuha, N., & Yusuf, A. (2021). Model Prediksi Kelulusan Tepat Waktu Dengan Metode Fuzzy C-Means Dan K-Nearest Neighbors Menggunakan Data Registrasi Mahasiswa. *Network Engineering Research*

- Operation*, 6(1), 38–46.
- Otok, B. W., Rachmawati, D. S., Purnami, S. W., & Mahpolah, M. (2021). Poisson regression modeling of diarrhea events in pasuruan regency with maximum likelihood estimates and generalized method moment. *International Journal of Preventive Medicine*, 12.
- Otok, B. W., Suharsono, A., Purhadi, P., Standsyah, H. A. A., & Al Azies, H. (2022). Partitional Clustering of Underdeveloped Area Infrastructure with Unsupervised Learning Approach: A Case Study in the Island of Java, Indonesia. *Journal of Regional and City Planning*, 33(2), 29–48.
- Prasetya, A. A. (2021). SST., M. Ec. Dev. Tsuraya Mumtaz, SST. Novia Dwima Santi, S. Tr. *Stat.*
- Prokhorov, A. V, Gubin, M. Y., Shesterikov, A. V, Arsenin, A. V, Volkov, V. S., & Evlyukhin, A. B. (2023). Lasing Effect in Symmetrical van der Waals Heterostructured Metasurfaces Due to Lattice-Induced Multipole Coupling. *Nano Letters*.
- Purwandari, A., & Wahyuningtyas, D. T. (2017). Eksperimen model pembelajaran teams games tournament (tgt) berbantuan media keranjang biji-bijian terhadap hasil belajar materi perkalian dan pembagian siswa kelas ii sdn saptorenggo 02. *Jurnal Ilmiah Sekolah Dasar*, 1(3), 163–170.
- Soemartini, S., & Supartini, E. (2017). *Analisis K-Means cluster untuk pengelompokan kabupaten/kota di Jawa Barat berdasarkan indikator masyarakat*.
- Standsyah, R. E., Otok, B. W., & Suharsono, A. (2023). Estimation of poverty model effect size in Indonesia using meta-analytic structural equation modeling (MASEM). *AIP Conference Proceedings*, 2588(1).
- Talakua, M. W., Leleury, Z. A., & Taluta, A. W. (2017). Analisis cluster dengan menggunakan metode k-means untuk mengelompokkan Kabupaten/Kota di provinsi maluku berdasarkan indikator indeks pembangunan manusia tahun 2014. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 11(2), 119–128.
- Wei, Y.-M., Chen, K., Kang, J.-N., Chen, W., Wang, X.-Y., & Zhang, X. (2022). Policy and management of carbon peaking and carbon neutrality: A literature review. *Engineering*, 14, 52–63.
- Wu, B., Gu, Q., Liu, Z., & Liu, J. (2023). Clustered institutional investors, shared ESG preferences and low-carbon innovation in family firm. *Technological Forecasting and Social Change*, 194, 122676.
- Xiao, K., Lin, Y.-H., Zhang, M., Oliver, R. D. J., Wang, X., Liu, Z., Luo, X., Li, J., Lai, D., & Luo, H. (2022). Scalable processing for realizing 21.7%-efficient all-perovskite tandem solar modules. *Science*, 376(6594), 762–767.
- Yu, W., & Sepehrnoori, K. (2013). Simulation of gas desorption and geomechanics effects for unconventional gas reservoirs. *SPE Western Regional & AAPG Pacific Section Meeting 2013 Joint Technical Conference*.
- Zhu, B., & Stülke, J. (2018). Subti Wiki in 2018: from genes and proteins to functional network annotation of the model organism *Bacillus subtilis*. *Nucleic Acids Research*, 46(D1), D743–D748.