Jurnal Riset Sains dan Teknologi
Volume 6 No. 1 Maret 2022, 61-64

e-ISSN: 2549-9750
p-ISSN: 2579-9118

Application of the K-Means Cluster for the Classification of Disadvantaged Districts/Cities in Maluku Province

Muhammad Yahya Matdoan[1*], Faraniena Yunaeni Risdiana[2], Gabriella Haumahu[3]
[1,3]Statistics Study Program, FMIPA, University of Pattimura, Ambon, Indonesia
[2]Statistics Study Program, State Islamic Institute Madura, Indonesia

*email: keepyahya@gmail.com

DOI; 10.30595/jrst.v6i1.11637

Histori Artikel:

Diajukan: 02/09/2021

Diterima: 02/11/2022

Diterbitkan: 11/11/2022

ABSTRAK

Maluku Province is still the 4th poorest province in Indonesia. This is due to the disparity in development between the provincial and district centers, cities and villages as well as government work programs that are not implemented evenly. To overcome and evaluate these problems, it is necessary to plan or study the classification of underdeveloped regions, namely by grouping districts/cities based on indicators of nderdeveloped areas. This research was conducted using secondary data obtained from the Central Statistics Agency (BPS) of Maluku Province. The method used in this study is to use the K-Means Cluster analysis method. The results of the study indicate that there are 2 classifications of underdeveloped and undeveloped areas in Maluku Province. Cluster 1 consists of Tanimbar Islands Regency, Southeast Maluku Regency, Central Maluku Regency, Buru Regency, Aru Islands Regency, West Seram Regency, Eastern Seram Regency, Southwest Maluku Regency, South Buru Regency and Tual City. In Cluster 2 there is only one area, namely Ambon City.

Kata Kunci: Classification, K-Means, Disadventages Areas, Maluku

1. INTRODUCTION

Since June 17, 1958, Maluku has been designated by the government as a province in Indonesia, following the provinces of Sumatra, West Java, Central Java, East Java, Lesser Sunda, Borneo and Sulawesi. This makes Maluku Province one of the oldest provinces in Indonesia. Based on the 2020 population census published by BPS Maluku Province, it shows that the population in Maluku Province is 1,831,880 people. Until now, the number of regencies/cities in Maluku Province is 11 regencies/cities consisting of 9 regencies and 2 cities. In addition, Maluku Province has 1,412 islands, including large and small islands, with a coastline length of 10,662 km with an area distribution of 92.4% ocean and 7.6% land so that Maluku Province can be said to be an Archipelago Province.

Based on data from the Central Statistics Agency (BPS) in 2020, it shows that Maluku Province is still ranked the 4th poorest province in Indonesia after Papua, West Papua and East Nusa Tenggara (NTT) with the percentage of poor people in Maluku Province of 17.99%. This is due to the disparity in development between the provincial and district centers, urban and rural areas and the programs planned and implemented by the government are uneven and not well targeted. In addition, the development budget for the regions is inadequate with the geographical location of the islands that are difficult to reach, which is the main cause of regional development in Maluku Province. The following data is presented in Table 1, the distribution of the percentage of poor people in Maluku Province in 2020.

Table 1. Distribution of the Percentage of Poor Population in Maluku Province

County/City	Percentage of Poor Population
Kepulauan Tanimbar	27,22
Maluku Tenggara	22,57
Maluku Tengah	19,83
Buru	16,64
Kepulauan Aru	26.26
Seram Bagian Barat	25,11
Seram Bagian Timur	23,04
Maluku Barat Daya	29,15
Buru Selatan	15,75
Ambon	4,51
Tual	22,51

Based on Table 1, it can be seen that the district with the highest percentage of poor people is Southwest Maluku (MBD) at 29.15%, then the Tanimbar Islands Regency is 27.22% and the Aru Islands Regency is 26.26%. Furthermore, the districts/cities that have the lowest percentage of poor people are Ambon City with a percentage of poor people at 4.51%, then South Buru District at 15.75% and Buru District with a percentage of 16.64% (BPS Maluku Province, 2021). In addition, data from the Ministry of Finance shows that of the 11 regencies/cities in Maluku Province, there are 8 regencies that are categorized as underdeveloped areas and experience inequality (Kementerian Keuangan, 2021). This is reflected in the high disparity between regions in terms of education, economy, infrastructure, and the quality of human resources.

Figure 1. Map of Maluku Province

Underdeveloped districts are districts where the community and the area they live in are relatively underdeveloped based on the categories of community economy, infrastructure, human resources (HR), regional financial capacity, accessibility, and regional characteristics (Mustika adn Pujiono, 2017). To overcome and evaluate these problems, it is necessary to have a system or program that can classify disadvantaged districts/cities in Maluku Province. One of the methods used in the classification is the cluster analysis method.

Cluster analysis is a multivariate technique that has the main goal of grouping objects based on their characteristics (Poerwanto, B., & Fa'rifah, R. Y, 2016). Cluster analysis classifies objects so that each object with the closest similarity to another object is in the same cluster. The clusters formed have high internal homogeneity and high external heterogeneity. In contrast to other multivariate techniques, this analysis does not estimate the set of variables empirically instead using the set of variables determined by the researcher himself (Suhaeni, C., Kurnia, A., & Ristiyanti, R, 2018). Meanwhile, K-Means Cluster Analysis is a non-hierarchical cluster analysis method that seeks to partition existing objects into one or more clusters or groups of objects based on their characteristics, so that objects that have the same characteristics are grouped in the same cluster and objects that have the same characteristics. different groups are grouped into other clusters (Ningrat, D. R, 2016). The advantages of the K-Means method include the fact that this algorithm has a fairly high accuracy of object size, so this algorithm is relatively more scalable and efficient for processing large numbers of objects. In addition, the K-Means algorithm is not affected by the order of objects (Simamora, 2005).

2. RESEARCH METHODS

2.1 Data Source

The research data comes from econdary data obtained from the Central tatistics Agency (BPS) Maluku Province. he research objects are 11 districts / cities in Maluku Province.

2.2 Research Variable

The research variables used in this study are based on a number of indicators for underdeveloped areas. The indicators measured are:

a. Percentage of Poor Population.
b. Number of villages/wards that have school facilities.
c. Life Expectancy.
d. Number of Puskesmas.
e. Percentage of Households Using Electricity.
f. Number of Junior High Schools

2.3 Method of Analysis

The analysis method used in this research is using cluster k-means analysis with the following stages[5]:

a. Partition the objects under study into the initial K cluster.
b. Specifies the object that has the closest distance to the centroid (mean). (The distance measure used is Euclidean distance, used for both standardized and non-standardized observations). Recalculate the centroid that the cluster uses to obtain new objects and for clusters that are missing objects.
c. Repeat 2nd step until no more objects are moved. Euclidean distance between $\boldsymbol{x} = [x_1, x_2, \dots, x_p]^T$ and $\boldsymbol{y} = [y_1, y_2, \dots, y_p]^T$

that is:

$$d(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_p - y_p)^2}$$

3. RESULT AND DISCUSSION

Based on the results of data processing, the results obtained can be seen in Table 1, the initial cluster shows that in the early stages 2 clusters were formed. Then the K-Means Cluster method will test and iterate for data recolation until it is found that there are no more objects moving clusters.

Table 2. *Initial Cluster Centers*

	Cluster	
	1	**2**
X_1	.74221	-2.40738
X_2	.35545	-1.19135
X_3	-.55612	1.99277
X_4	1.11844	-.14989
X_5	-2.22925	1.03014
X_6	-.61913	-.38429

Based on Table 2, the final result is obtained Cluster Center, contains 2 clusters that divide district/city data in Maluku Province based on 6 variables, namely Percentage of Poor Population (X_1), Number of Villages that have school facilities (X_2), Life Expectancy Rate (X_3), Number of Puskesmas (X_4), Percentage of Household Users Electricity (X_5) and Number of Junior High Schools (X_6). The distance of each cluster to other clusters.

Table 3. Distances between Final Cluster Centers

Cluster	**1**	**2**
1		3.876
2	3.876	

The distance between Cluster 1 and Cluster 2 is very close, and have things in common close when compared to a while cluster 2 The number of members / data per cluster shown in Table 4.

Table 4. Cluster Membership

District/City	**Cluster**	**Distance**
Kepulauan Tanimbar	1	1.301
Maluku Tenggara	1	1.145
Maluku Tengah	1	3.758
Buru	1	1.907
Kepulauan Aru	1	2.570
Seram Bagian Barat	1	1.532
Seram Bagian Timur	1	1.812
Maluku Barat Daya	1	2.128
Buru Selatan	1	1.813
Ambon	2	.000
Tual	1	2.351

From Table 4, it can be clustered regencies/cities in Maluku Province is based on the indicators of underdeveloped areas, namely cluster 1 consisting of Kepulauan Tanimbar Regency, Maluku Tenggara Regency, Maluku Tengah Regency, Buru Regency, Kepualauan Aru Regency, Seram Bagian Barat Rgency, Seram Bagian Timur, Maluku Barat Daya Regency, Buru Selatan Regency and the City ual. In cluster 2, there is only one area, namely Ambon City.

4. CONCLUSION

Based on the results and discussion, it is concluded that the classification of underdeveloped areas in districts / cities in Maluku Province using the K-Means method is cluster 1 consisting of Kepeulauan Tanimbar Regency, Maluku Tenggara Regency, Maluku Tengah Regency, Buru Regency, Kepualauan Aru Regency, Seram Bagian Barat Regency, Seram Bagian Timur Regency, Maluku Barat Daya Regency, Buru Selatan and Tual City. In cluster 2, there is only one area, namely Ambon City.

REFERENCES

Indonesia, K. M. K. R., & BEA, P. H. E. U. P. (2021). Kementerian Keuangan Republik Indonesia.

Maluku, B. P. S. (2021). Provinsi Maluku Dalam Angka 2017. *Ambon: BPS Maluku*.

Mustika, D., & Pujiyono, A. (2017). *Pengaruh Infrastruktur terhadap PDRB Kabupaten tertinggal di provinsi Nusa Tenggara Timur* (Doctoral dissertation, Fakultas Ekonomika dan Bisnis).

Ningrat, D. R. (2016). *Analisis Cluster Dengan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi* (Doctoral dissertation, Fakultas Sains dan Matematika, Undip).

Poerwanto, B., & Fa'rifah, R. Y. (2016). Analisis cluster k-means dalam pengelompokan kemampuan mahasiswa. *Indonesian Journal of Fundamental Sciences*, *2*(2), 92-96.

Simamora, B. (2005). *Analisis multivariat pemasaran*. Gramedia Pustaka Utama.

Suhaeni, C., Kurnia, A., & Ristiyanti, R. (2018). Perbandingan Hasil Pengelompokan menggunakan Analisis Cluster Berhirarki, K-Means Cluster, dan Cluster Ensemble (Studi Kasus Data Indikator Pelayanan Kesehatan Ibu Hamil). *Jurnal Media Infotama*, *14*(1).