

COMPARISON OF REGIONAL CLUSTER ANALYSIS ACCORDING TO INCLUSIVE DEVELOPMENT INDICATORS IN JAVA ISLAND 2018 BETWEEN HIERARCHICAL AND PARTITIONING CLUSTERING STRATEGIES

Akhmad Fatikhurrizqi¹; Arie Wahyu Wijayanto²

Statistika^{1,2}

Politeknik Statistika STIS

www.stis.ac.id

211709518@stis.ac.id¹; ariewahyu@stis.ac.id²

Abstract— Gross Domestic Product (GDP) is one of the most common indicators to reflect a nation's development. Indonesia's GDP has an average growth rate of 5 percent over the 2015-2019 period with the highest growth rate occurred in 2018. Furthermore, the provinces in Java Island contributed the most out of any province to Indonesia's GDP in that year. However, the development in Java Island still has several issues, such as high poverty, unequal income distribution, and high unemployment. This problem indicates that the economic growth in Java Island has not been inclusive concerning development. This study aims to group regencies/municipalities in Java Island based on indicators of inclusive growth. These indicators refer to McKinley (2010) in a journal published by the Asian Development Bank (ADB). The cluster methods used to represent each hierarchical and partitioning are the Agglomerative Nesting (AGNES) and K-Means methods. The results of this study show that there are 3 clusters based on the AGNES method and 4 clusters based on the K-Means method. Clusters with good inclusive growth characteristics are dominated by municipality areas based on the K-Means method. Meanwhile, the clusters with low inclusive growth characteristics are dominated by regencies/municipalities on Madura Island based on the K-Means and AGNES methods. The comparison of the appropriate methods in this study based on the silhouette value is the AGNES method.

Keywords: Inclusive Growth, Cluster Analysis, AGNES, K-Means.

Abstrak— Produk Domestik Bruto (PDB) adalah salah satu indikator yang paling umum untuk mencerminkan pembangunan suatu bangsa. PDB Indonesia memiliki tingkat pertumbuhan rata-rata 5 persen selama periode 2015-2019 dengan laju pertumbuhan tertinggi terjadi pada tahun 2018. Selain itu, provinsi-provinsi di Pulau Jawa memberikan kontribusi terbesar dibanding provinsi lainnya terhadap PDB Indonesia pada tahun tersebut. Namun demikian, pembangunan di Pulau Jawa masih memiliki beberapa permasalahan, seperti kemiskinan yang tinggi, distribusi pendapatan yang tidak merata, dan pengangguran yang tinggi. Masalah ini menunjukkan bahwa pertumbuhan ekonomi di Pulau Jawa belum inklusif dalam pembangunan. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Pulau Jawa berdasarkan indikator pertumbuhan inklusif. Indikator tersebut mengacu pada McKinley (2010) dalam jurnal yang diterbitkan oleh Asian Development Bank (ADB). Metode klaster yang digunakan untuk merepresentasikan setiap hierarki dan partisi adalah metode Agglomerative Nesting (AGNES) dan K-Means. Hasil penelitian menunjukkan bahwa terdapat 3 klaster berdasarkan metode AGNES dan 4 cluster berdasarkan metode K-Means. Klaster dengan karakteristik pertumbuhan inklusif yang baik didominasi oleh kawasan perkotaan berdasarkan metode K-Means. Sedangkan klaster dengan karakteristik pertumbuhan inklusif rendah didominasi oleh kabupaten/kota di Pulau Madura berdasarkan metode K-Means dan AGNES. Perbandingan metode yang sesuai dalam penelitian ini berdasarkan nilai silhouette adalah dengan metode AGNES.

Kata Kunci: Pertumbuhan Inklusif, Analisis Klaster, AGNES, K-Means.

INTRODUCTION

The concept of inclusive growth is interpreted narrowly and broadly [1]. In a narrow sense, inclusive growth focuses more on economic growth, where human capabilities are seen as an

instrument to increase economic output. Whereas in a broad sense, inclusive growth more focuses on comprehensive and sustainable growth, which will create and expand economic opportunities, and ensure wider access to these opportunities, so that every member of society can participate and feel



the benefit from this growth. This shows that economic growth, which is commonly used as a measure of a country's development, only includes a narrow concept, where the measure of economic growth is emphasized on an increase in output or Gross Domestic Product [2]. *Badan Pusat Statistik* (BPS) noted that Indonesia's economic growth rate in 2015-2019 experienced relatively good and quite stable with an average growth of around 5 percent, where the highest GDP growth rate occurred in 2018 [3]. Meanwhile, based on the regional contribution to Indonesia's GDP, Java Island contributed more than half of it. The contribution of Java Island to Indonesia's GDP in 2018 was 58.48 percent [3]. This is also supported by the large number of industrial estates in Java, which are the largest in number compared to other islands. Java Island is considered more ready to become a location for industrial development in terms of human resources, infrastructure, and natural resources [4]. Therefore, it can be said that Indonesia's economic growth is largely supported by the island of Java, even though it only has six provinces. The amount of Java Island's GRDP contribution to national GDP still leaves several problems, such as high poverty, unequal income

distribution, and high unemployment. In Table 1 it can be seen that all provinces in Java Island have relatively high economic growth, which is above the national economic growth value of only 5.17 percent in 2018 [3]. However, not all provinces have achieved high economic growth make a good contribution to public welfare. The emphasis on increasing GDP leads to a subordination of problems such as poverty, discrimination, unemployment, and income distribution [2]. Also, GDP per capita as a measure of economic growth is marred by issues of inequality and severe poverty. Therefore, inclusive growth requires the growth and improvement of individual welfare, especially in the income and non-income aspects of individual and societal well-being [5]. This condition goes along with the condition of several problems in Java Island, such as the percentage of poverty in Central Java, DI Yogyakarta, and East Java having values above 10 percent and exceeding the national average; the open unemployment rate in DKI Jakarta, West Java, and Banten also scored above the national average [6]; as well as the distribution of income inequality in DKI Jakarta, West Java, and DI Yogyakarta also scored above the national average.

Table 1. The value of economic growth, the percentage of poverty, the open unemployment rate, and the Gini ratio of provinces in Java Island in 2018

Province	Economic Growth	Percentage of Poverty	Unemployment Rate	Gini Ratio
DKI Jakarta	6.17	3.55	6.24	0.39
Jawa Barat	5.66	7.25	8.17	0.41
Jawa Tengah	5.31	11.19	4.51	0.36
DI Yogyakarta	6.20	11.81	3.35	0.42
Jawa Timur	5.50	10.85	3.99	0.37
Banten	5.82	5.25	8.52	0.37
Nasional	5.17	9.66	5.34	0.38

Source: BPS [7]

Based on Tabel 1, Conditions of poverty, unemployment, and high inequality of income distribution indicate that economic growth in Java is not yet inclusive. Inclusive growth is growth followed by a reduction in poverty, a reduction in inequality in income distribution, and an increase in labor absorption [8]. Inclusive growth reflects more on regional growth in a multidimensional manner, so that good linkages between indicators can support inclusive growth. For example, income inequality affects inclusive education [9]. Early childhood education is an effective target for fiscal spending to create employment and gender-inclusive growth [10]. Also, case studies in Africa show the health sector to play a major role in inclusive growth [11].

Even though it isn't a new issue, inclusive growth is still a topic that attracts attention and is

considered important because it can inclusively explain the state of development, so that it is useful for promoting sustainable development. Therefore, this study discusses the grouping of regencies/municipalities in Java based on their level of growth inclusiveness. The objectives of this study are as follows:

- 1) Grouping regencies/municipalities in Java Island based on indicators of inclusive growth based on the cluster analysis method [12].
- 2) Knowing the characteristics of regencies/municipalities in Java Island based on the clusters formed.
- 3) Comparing between Agglomerative Nesting (AGNES) and K-Means clustering methods as representatives of each hierarchical and partitioning strategy.

MATERIALS AND METHODS

Data Source

The data source of this study is the Central Bureau of Statistics (BPS). Among other things, the publication of each province in Java, such as “*Provinsi Dalam Angka 2019*”, “*Statistik Kesejahteraan Rakyat Provinsi 2018*”, dan “*Keadaan Angkatan Kerja Provinsi 2018*”, as well as the publication of tables and indicators from the website of the BPS. The units of analysis used in this study were 119 regencies/municipalities in Java Island in 2018.

Research Variable

The relatively new concept of inclusive growth makes it difficult to determine a theory that supports the argument for the direction of the influence of the variables used in describing inclusive growth [13]. This study uses the concept of the Asian Development Bank which explains the dimensions of inclusive growth [1]. The variables used in this study which refer to these dimensions can be seen in

Table 2 below.

Table 2. Research Variable

Dimension	Variable	Definition	Unit of Measurement
Economic Growth	X1	Grow rate of GRDP per capita	Percent
Productive Workforce	X2	Percentage of Labor in the Industrial Sector	Percent
	X3	Unemployment rate (TPT)	Percent
Economic Infrastructure	X4	Percentage of households with access to electricity	Percent
	X5	Percentage of households that own a computer/laptop	Percent
Poverty	X6	Percentage of the poor population	Percent
Inequality	X7	Gini Ratio	Ratio
	X8	Gender Development Index (IPG)	Ratio
Gender Equality	X9	Gender Empowerment Index (IDG)	Ratio
	X10	Life expectancy at birth (UHH)	Year
Health	X11	The expectation of old school years	Percent
Education	X12	Percentage of households that have access to proper sanitation	Year
	X13	Percentage of households that have access to decent drinking water	Percent
Water and Sanitation	X14	Percentage of households buying/receiving “Raskin”	Percent
	X15	Percentage of the population having social protection card (KPS)	Percent

Source: McKinley [1], modified

Method of Analysis

1) **AGNES (Agglomerative Hierarchical Clustering)**

An agglomerative hierarchical clustering method uses a bottom-up strategy. It typically starts by letting each object form its cluster and iteratively merges clusters into larger and larger clusters, until all the objects are in a single cluster or certain termination conditions are satisfied. The single cluster becomes the hierarchy’s root. For the merging step, it finds the two clusters that are closest to each other (according to some similarity measure) and combines the two to form one cluster. Because two clusters are merged per iteration, where each cluster contains at least one object, an agglomerative method requires at most n iterations [14]. How AGNES works can be seen in Figure 1 below.

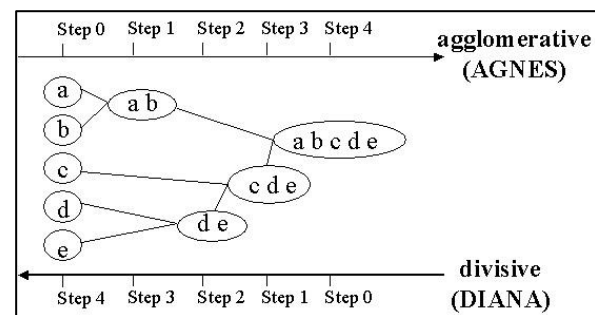


Figure 1. Concept of Hierarchical Clustering

2) **K-means**

K-Means Clustering is an unsupervised clustering algorithm. This method groups several objects into a cluster. Membership of objects is seen based on the object's distance to the center point (centroid) of a cluster. In the k-means approach, the initial step is to determine k as the

number of clusters to be formed. The k-means algorithm can be written as follows [15]:

- a) For n objects, initiate k the cluster center point
- b) Enter each object in the closest cluster, calculating the distance.

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \dots\dots\dots (1)$$

- c) Update the cluster center point for each incoming object, looking at the average member.
- d) Repeat steps 2 and 3 until there is no change in the center point of the cluster.

The systematics and workings of the k-means algorithm can be seen in Figure 2 below.

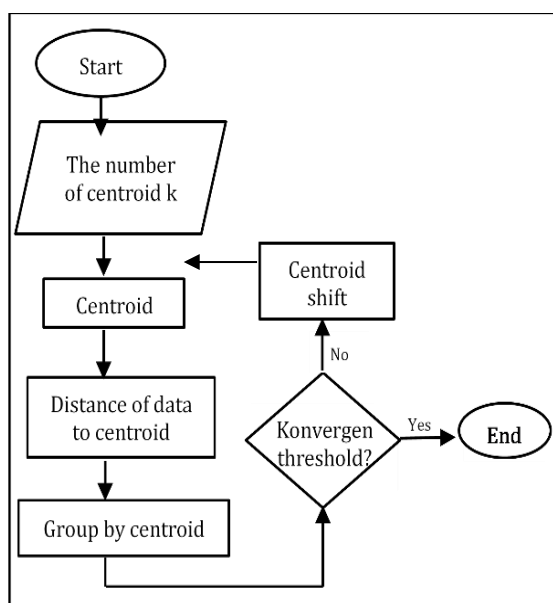


Figure 2. Flowchart of K-Means Algorithm

RESULTS AND DISCUSSION

Pre-processing Data

Missing Value

Based on the available data, there is a missing value in the Gini ratio variable in the regencies/municipalities of DKI Jakarta Province. Therefore, steps that can be taken to overcome this problem are to estimate the Gini ratio using the linear regression method. Modeling was carried out on other regencies/municipalities in Java Island. The independent variables used are per capita expenditure (pkp) and the level of poverty depth (p1). The model equation that is formed is as follows:

$$Gini\ ratio_i = 0,276 + (6,551 \times 10^{-6})pkp_i - 0,011 p1_i \dots\dots\dots (2)$$

Standardization / Normalization

Normalization is a step for adjusting the data values into specific ranges such as between 0 to 1

or -1 to 1 [16]. The purpose of normalization is to equalize the units in all variables. Generally, the methods used are minimum-maximum, z-score, and decimal scaling. In this study, the z-score normalization was used.

Feature Selection

Feature selection is the stage of selecting the variables to be used. This stage can reduce the number of variables, get rid of unrelated, repetitive, or noise variables, thereby speeding up the data mining algorithm, increasing accuracy, and producing the best model [17]. In this study, the variables were selected based on the assumption of non-multicollinearity. A low correlation value between variables indicates the absence of data redundancy. Therefore, a variable that has a high correlation value is selected, which is above 0.7. Figure 3 shows the pairs of variables that correlate 0.7 are X5-X6, X6-X14, X6-X15, and X14-X15. The decision was taken, namely to issue X14 and X15 because it was already represented by X6, and to issue X5 because other variables represented the dimensions of economic infrastructure.

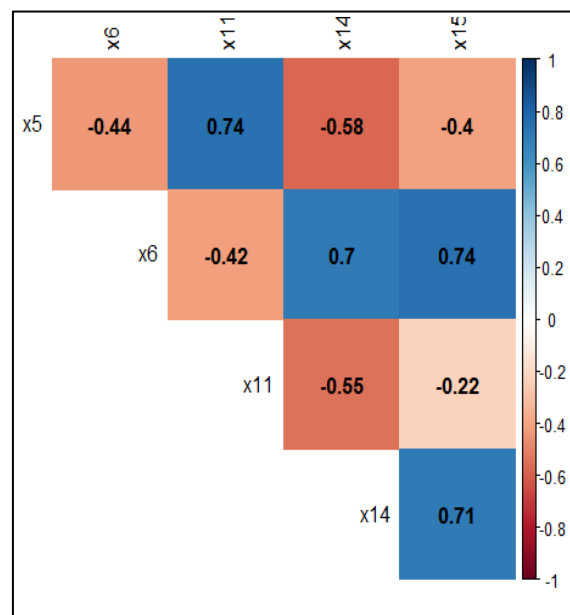


Figure 3. Correlation of X5, X6, X11, X14, X15

Result of the AGNES (Agglomerative Nesting) Clustering Analysis

The AGNES method is a hierarchical cluster method that moves from bottom to top. As for the hierarchical method, there are several methods such as Single Linkage, Average Linkage, Ward, and Complete. In determining the most appropriate method, the Agglomerative Coefficient can be used. The coefficient value close to 1 indicates the grouping is getting stronger or better. The

following is the Agglomerative Coefficient value for each method:

Method	Agglomerative Coefficient
Average	0,7541
Single Linkage	0,6534
Complete Linkage	0,8088
Ward	0,8838

Based on Table 3, the best method of hierarchical grouping in this study is the ward method, with an agglomerative coefficient value of 0.88. In the AGNES hierarchy method, it can be seen that the cluster formation is based on the formed dendrogram. Figure 4 below is the dendrogram that is formed.

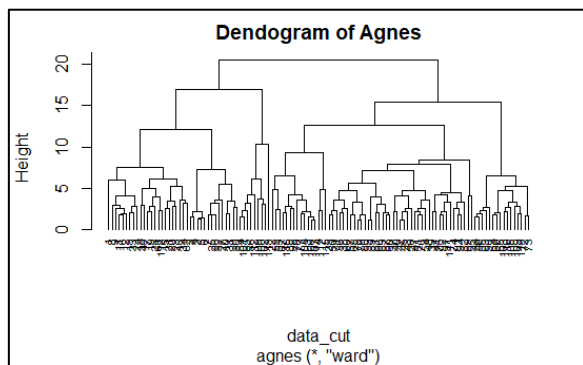


Figure 4. Dendrogram of Grouping with AGNES Method

Furthermore, in determining the maximum number of clusters to be formed based on the AGNES method with the ward, the NbClust () package is used in the R application. The results obtained show that the best number of clusters is three clusters. The number of members in each cluster formed is 37, 72, and 9 regencies/municipalities. As for the distribution of members in each cluster, it can be seen in Figure 5 below.

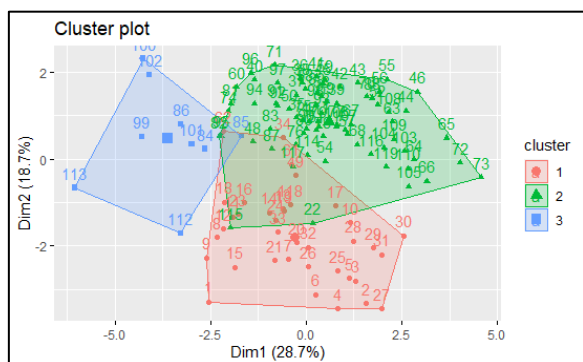


Figure 5. Clustering Plot based AGNES Method

Based on the average value of the variables in each cluster (Table 4), the characteristics of each

variable in each cluster are obtained, which are as follows:

1. Cluster 1 has unfavorable characteristics in several variables, namely the high level of open unemployment and the low percentage of households that have access to proper drinking water.
2. Cluster 2 has good characteristics in all variables, except poverty. This is indicated by the average percentage of poor people in cluster 2 of 9.76 percent, which is still above the national average (9.66 percent).
3. Cluster 3 has poor characteristics in several variables, namely the per capita GRDP growth rate and the lowest percentage of industrial workers among other clusters, the high average percentage of poor people, and the low percentage of households with access to proper sanitation.

In general, all the clusters that were formed were still experiencing their problems, so that no cluster had good inclusive growth. Therefore, if sorted according to the level of inclusive growth from the highest are cluster 2, cluster 1, and cluster 3. The categories for each cluster are based on their inclusive growth, namely cluster 2 is quite good, cluster 1 is medium, and cluster 3 is not good.

Table 4. Average of Variable Value Based on AGNES Clustering

No	Variable	Cluster 1	Cluster 2	Cluster 3
1	X1	8.69	7.81	7.77
2	X2	18.14	25.16	17.62
3	X3	7.42	4.42	4.26
4	X4	99.96	99.9	99.67
5	X6	7.94	9.76	15.37
6	X7	0.36	0.33	0.28
7	X8	89.58	92.14	84.83
8	X9	66.78	69.32	58.58
9	X10	71.86	73.27	67.63
10	X11	12.70	13.23	12.6
11	X12	64.10	84.53	33.45
12	X13	36.13	76.84	62.94

The results of grouping regencies/municipalities based on inclusive growth clusters using the AGNES method can be seen in Table 5. Cluster 1 is dominated by regencies/municipalities in DKI Jakarta and West Java. Cluster 2 is dominated by regencies/municipalities in Central Java and DI Yogyakarta. The whole cluster 3 is a regency area and is dominated by areas on Madura Island. The results of grouping using the AGNES method show the tendency of regencies/municipalities within a

province horde in a cluster. This indicates there is a spatial influence between regions in explaining the inclusiveness of a region's growth [18].

Table 5. Results of Grouping of Regencies/Municipalities with AGNES Method

Cluster	Regencies/Municipalities
Cluster 1	Kep. Seribu, Kota Jakarta Selatan, Kota Jakarta Timur, Kota Jakarta Pusat, Kota Jakarta Barat, Kota Jakarta Utara, Bogor, Sukabumi, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar, Cilacap, Blora, Tegal, Brebes, Kota Serang.
Cluster 2	Bekasi, Banyumas, Purbalingga, Banjarnegara, Kebumen, Purworejo, Wonosobo, Magelang, Boyolali, Klaten, Sukoharjo, Wonogiri, Karanganyar, Sragen, Grobogan, Rembang, Pati, Kudus, Jepara, Demak, Semarang, Temanggung, Kendal, Batang, Pekalongan, Pemalang, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, Kota Tegal, Kulon Progo, Bantul, Gunung Kidul, Sleman, Kota Yogyakarta, Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Malang, Lumajang, Jember, Banyuwangi, Pasuruan, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Gresik, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu, Tangerang, Serang, Kota Tangerang, Kota Cilegon, Kota Tangerang Selatan.
Cluster 3	Bondowoso, Situbondo, Probolinggo, Bangkalan, Sampang, Pamekasan, Sumenep, Pandeglang, Lebak

Results of the K-means Cluster Analysis

The initial stage of k-means is to determine the optimum number of clusters. The method used in determining the number of clusters in this study is the Elbow Method. The smaller the variance in the cluster (within), the better the grouping is done. Figure 6 shows an elbow plot where the horizontal axis is the number of clusters and the vertical axis is the total number of squares in the cluster (total within the sum of squares). Determining the optimum number of clusters is by determining the point that has decreased sharply from the total value within the sum of the square, where this value can determine the variance value in the cluster.

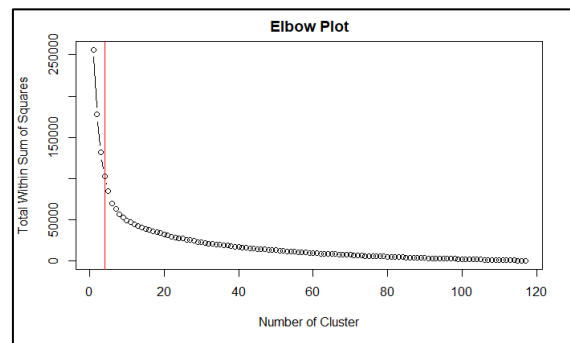


Figure 6. Elbow Plot

Based on these results, it illustrates that the number of clusters is 4 can significantly reduce the within variance, while the number of clusters more than 4 can only reduce the within variance which is less, so it is less efficient. Also, Figure 7 shows that the cluster formed can separate objects well and there is no overlapping.

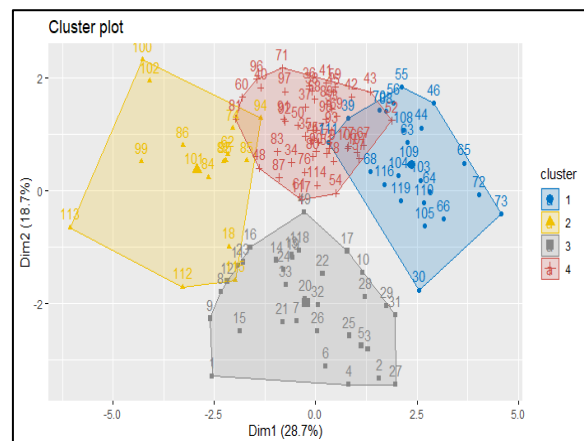


Figure 7. Clustering Plot based K-Means Method

The results of grouping using the K-means method indicate that there are 4 clusters formed, each of which has members of 24, 16, 33, and 46 regencies/municipalities. The variance between clusters formed was 516.8 and the variance within each cluster was 143.8; 199.2; 267.8; 288.4. This shows that the cluster formed already has good cluster characteristics, namely variance between high clusters (heterogeneous) and variance in low clusters (homogeneous). Based on the average value of the variables in each cluster (Table 6), the characteristics of each variable in each cluster are obtained, which are as follows:

1. Cluster 1 has good characteristics in all variables, meaning that it can be said that the regencies/municipalities in Cluster 1 already have inclusive development.
2. Cluster 2 has unfavorable characteristics in several variables, namely the lowest percentage of workers in the industrial sector among other clusters, high poverty, low levels

- of gender equality, and a low percentage of households with access to proper sanitation.
- Cluster 3 has unfavorable characteristics in several variables, namely the high level of open unemployment and the low percentage of households that have access to proper drinking water.
 - Cluster 4 has fairly good characteristics in all dimensions, except poverty. This is indicated by the average percentage of poor people of 11.15 percent which has a value above the national average (9.66 percent).

Therefore, if all clusters sorted by the level of inclusive growth from the highest, there are cluster 1, cluster 4, cluster 3, and cluster 2. The categories for each cluster are based on inclusive growth, namely cluster 1 is good, cluster 4 is good enough, cluster 3 is moderate, and cluster 2 is not good.

Table 6. Average of Variable Value Based on K-Means Clustering

Variable	Cluster 1	Cluster 2	Cluster 3	Cluster 4
X1	7.72	8.02	8.84	7.75
X2	22.77	18.24	19.20	25.97
X3	4.53	5.02	7.45	4.35
X4	99.92	99.66	99.97	99.93
X6	6.68	13.99	7.51	11.15
X7	0.35	0.29	0.37	0.32
X8	94.90	86.25	89.76	90.98
X9	76.12	61.40	66.45	66.44
X10	74.56	68.50	71.93	72.98
X11	14.36	12.56	12.73	12.69
X12	93.10	47.31	61.66	82.98
X13	77.95	63.36	32.10	77.58

The results of grouping regencies/municipalities based on inclusive growth clusters using the K-means method can be seen in Table 7. Cluster 1 is dominated by municipality areas and several regencies that are quite advanced, such as Semarang and Sleman. In Cluster 2, all of its members are regencies, even all regencies on Madura Island are included in this cluster. Cluster 3 has a fairly even distribution between regency and municipality areas. This cluster also occurs in the grouping of all areas in DKI Jakarta. Finally, cluster 4 is dominated by regencies in Central Java and East Java.

Table 7. Results of Grouping of Regencies/Municipalities with K-means Method

Cluster	Regencies / Municipalities
Cluster 1	Kota Depok, Purworejo, Sukoharjo, Karanganyar, Semarang, Temanggung, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota

	Semarang, Kota Tegal, Bantul, Sleman, Kota Yogyakarta, Sidoarjo, Kota Kediri, Kota Blitar, Kota Malang, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu, Kota Tangerang, Kota Tangerang Selatan
Cluster 2	Indramayu, Brebes, Pacitan, Jember, Bondowoso, Situbondo, Probolinggo, Ngawi, Bojonegoro, Bangkalan, Sampang, Pamekasan, Sumenep, Pandeglang, Lebak, Serang.
Cluster 3	Kep. Seribu, Kota Jakarta Selatan, Kota Jakarta Timur, Kota Jakarta Pusat, Kota Jakarta Barat, Kota Jakarta Utara, Bogor, Sukabumi, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Cimahi, Kota Tasikmalaya, Kota Banjar, Blora, Kota Serang.
Cluster 4	Cilacap, Banyumas, Purbalingga, Banjarnegara, Kebumen, Wonosobo, Magelang, Boyolali, Klaten, Wonogiri, Sragen, Grobogan, Rembang, Pati, Kudus, Jepara, Demak, Kendal, Batang, Pekalongan, Pemalang, Tegal, Kota Pekalongan, Kulon Progo, Gunung Kidul, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Malang, Lumajang, Banyuwangi, Pasuruan, Mojokerto, Jombang, Nganjuk Madiun, Magetan, Tuban, Lamongan, Gresik, Kota Probolinggo, Kota Pasuruan, Tangerang, Kota Cilegon.

Comparison of The Result Clustering Analysis between AGNES and K-Means method

To compare the cluster method, this study used silhouette values. The results of the silhouette coefficient calculation can vary from -1 to 1. The coefficient of an object is 1 indicating that the object is in the right cluster. If it is 0, it indicates that the object is not certain which cluster it belongs to. Meanwhile, if the value is -1, it indicates that the cluster structure that is formed is overlapping.

Table 8. Silhouette Value

Method	Silhouette Value
AGNES	0,17
K-Means	0,15

Based on Table 8, the comparison of silhouette values between the AGNES and K-Means methods is not too different, namely 0.17 and 0.15. Apart from having a higher silhouette value, the AGNES method also has a smaller number of clusters. This shows that the AGNES method is more efficient in grouping objects. Therefore, it can be concluded that the AGNES method is more appropriate than the K-Means method in this study.

CONCLUSION

The grouping of regencies/municipalities in Java Island based on their level of inclusive growth resulted in several different clusters per method.

The AGNES method produces three clusters and the K-Means method produces four clusters. Clusters that have good inclusive growth characteristics are dominated by municipality areas based on the K-Means method. Meanwhile, regional clusters that have low inclusive growth based on the K-Means and AGNES methods are regencies/municipalities on Madura Island. The appropriate comparison method in this study between AGNES and K-means is based on silhouette values and efficiency in grouping regencies/municipalities in the AGNES method.

REFERENCE

- [1] T. McKinley, "Inclusive growth criteria and indicators: an inclusive growth index for diagnosis of country progress," *Asian Dev. Bank Work. Pap.*, no. 14, pp. 1–34, 2010, [Online]. Available: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Inclusive+Growth+Criteria+and+Indicators+:+An+Inclusive+Growth+Index+for+Diagnosis+of+Country+Progress#0>.
- [2] M. P. Todaro and S. C. Smith, *Economic Development Twelfth Edition*, 12th Editi. New York: Pearson, 2014.
- [3] Badan Pusat Statistik, "Data Online Badan Pusat Statistik: Tabel Dinamis Subjek Produk Domestik Bruto," 2019. <https://www.bps.go.id/subject/11/produk-domestik-bruto--lapangan-usaha-.html#subjekViewTab3>.
- [4] Kementerian Perindustrian RI, "Rencana Induk Pembangunan Industri Nasional 2015 - 2035," *Rencana Induk Pembang. Ind. Nas. 2015-2035*, pp. 1–98, 2015, [Online]. Available: <http://www.depkop.go.id>.
- [5] I. C. Oluseye and A. A. Gabriel, "Determinants of Inclusive Growth in Nigeria: An ARDL Approach," vol. 7, no. 3, pp. 97–109, 2017, DOI: 10.5923/j.economics.20170703.01.
- [6] T. Sukwika, "Peran Pembangunan Infrastruktur terhadap Ketimpangan Ekonomi Antarwilayah di Indonesia," *J. Wil. dan Lingkung.*, vol. 6, no. 2, p. 115, 2018, doi: 10.14710/jwl.6.2.115-130.
- [7] Badan Pusat Statistik, "Data Online Badan Pusat Statistik: Tabel Dinamis Subjek Produk Domestik Bruto, Kemiskinan dan Ketimpangan, serta Tenaga Kerja," 2020, [Online]. Available: <https://www.bps.go.id/subject/11/produk-domestik-bruto--lapangan-usaha-.html#subjekViewTab3>.
- [8] T. Ji Long and E. Pasaribu, "Analisis Spasial Determinan Pertumbuhan Inklusif Kabupaten/Kota Di Provinsi Jawa Tengah Tahun 2017," *Semin. Nas. Off. Stat.*, vol. 2019, no. 1, pp. 416–423, 2020, doi: 10.34123/semnasoffstat.v2019i1.11.
- [9] S. A. Asongu, J. Nnanna, and P. N. Acha-Anyi, "Finance, inequality and inclusive education in Sub-Saharan Africa," *Econ. Anal. Policy*, vol. 67, pp. 162–177, 2020, DOI: 10.1016/j.eap.2020.07.006.
- [10] K. Kim, İ. İlkkaracan, and T. Kaya, "Public investment in care services in Turkey: Promoting employment & gender-inclusive growth," *J. Policy Model.*, vol. 41, no. 6, pp. 1210–1229, 2019, DOI: 10.1016/j.jpolmod.2019.05.002.
- [11] S. Tella and O. Alimi, "Determinants of Inclusive Growth in Africa: Role of Health and Demographic Changes," *African J. Econ. Rev.*, vol. 4, no. 2, pp. 138–146, 2016.
- [12] S. Wahyuni and Y. A. Jatmiko, "Pengelompokan Kabupaten/Kota di Pulau Jawa Berdasarkan Faktor-Faktor Kemiskinan dengan Pendekatan Average Linkage Hierarchical Clustering," *J. Apl. Stat. Komputasi Stat.*, vol. 10, no. 1, pp. 1–8, 2019, doi: <https://doi.org/10.34123/jurnalasks.v10i1.197>.
- [13] N. Ngepah, "A review of theories and evidence of inclusive growth: an economic perspective for Africa," *Curr. Opin. Environ. Sustain.*, vol. 24, pp. 52–57, 2017, DOI: 10.1016/j.cosust.2017.01.008.
- [14] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Third Edit. Amsterdam: Morgan Kauffman, 2012.
- [15] S. K. Chandel, "Intrusion Detection System using K-Means Data Mining and Outlier Detection Approach," Masaryk University, 2017.
- [16] S. A. Alasadi and W. S. Bhaya, "Alasadi, S. A. (2017). Review of data preprocessing techniques in data mining 12(16), 4102–4107," *J. Eng. Appl. Sci.*, vol. 12, no. 16, pp. 4102–4107, 2017, DOI: <http://dx.doi.org/10.36478/jeasci.2017.4102.4107>.
- [17] H. Liu, H. Motoda, R. Setiono, and Z. Zhao, "Feature selection: An ever-evolving frontier in data mining," *Featur. Sel. data Min.*, vol. 10, pp. 4–13, 2010, [Online]. Available: <http://proceedings.mlr.press/v10/liu10b/liu10b.pdf>.
- [18] C. Park and R. Claveria, "DOES REGIONAL INTEGRATION MATTER FOR INCLUSIVE GROWTH? EVIDENCE FROM THE MULTIDIMENSIONAL REGIONAL ADB

ECONOMICS Does Regional Integration Matter for Inclusive Growth ? Evidence from the Multidimensional Regional Integration Index," *ADB Econ. Work. Pap. Ser.*, vol. Oktober, no. 559, 2018, [Online]. Available: <https://www.adb.org/sites/default/files/publication/460681/ewp-559-regional-integration-inclusive-growth.pdf>.

