

Grouping The Districts in Sumatera Region Based on Economic Development Indicators Using *K-Medoids* and CLARA Methods

Retsya Lapiza, Syafriandi*, Nonong Amalita, Dina Fitria

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: syafriandi_math@fmipa.unp.ac.id

Submitted : 16 Agustus 2022

Revised : 06 Desember 2022

Accepted : 14 Desember 2022

ABSTRACT

The regional regions of emerging countries frequently experience inequality in economic development. The districts of the Sumatera Region are one of the regional areas of Indonesia that have not yet had an equal distribution of economic development. The COVID-19 pandemic makes matters worse because it prevents communal activities and slows down economic growth. This study aims to identify regional groupings and evaluate the outcomes of those grouping. The K-Medoids and CLARA methods of cluster analysis were used in this study. K-Medoids and CLARA are useful for non-hierarchical cluster analysis for outliers. The data used in this study is secondary data sourced from the publication of the Badan Pusat Statistik (BPS) on Statistik Kesejahteraan Rakyat 2020. The analysis technique is carried out by examining outliers, normalizing the data, performing cluster analysis, and then validating the cluster results. The results of this study using the K-Medoids method showed that the first cluster had 59 districts and the second cluster had 95 districts. Meanwhile, the CLARA method with two groups produced clusters of 74 and 80 regions, respectively. The Silhouette Coefficient values using the K-Medoids and CLARA comparison methods are 0.13 and 0.15, respectively. As a result, in this study, the K-Medoids and CLARA methods gave grouping results with the same performance.

Keywords: CLARA, Economic Development, K-Medoids, Silhouette Coefficient, Sumatera Region

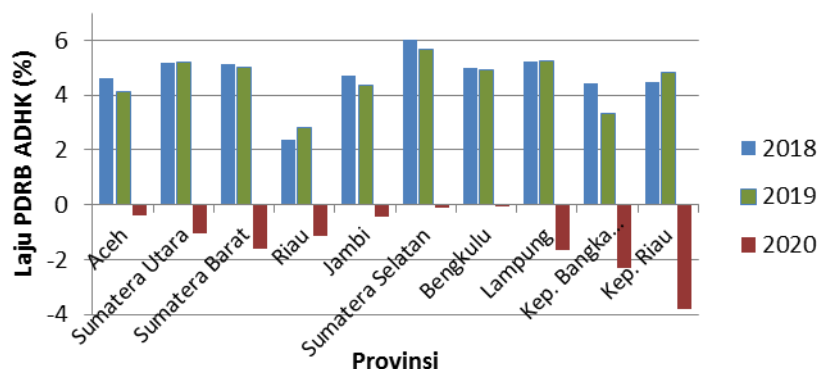


This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Pembangunan ekonomi adalah suatu proses yang menyebabkan kenaikan pendapatan riil perkapita penduduk disertai dengan perubahan aspek kelembagaan menjadi lebih baik dalam jangka waktu yang lama di suatu negara. Untuk mengetahui hasil pembangunan maka diperlukan indikator-indikator yang dapat menggambarkan pembangunan ekonomi suatu daerah. Pembangunan ekonomi suatu daerah diukur berdasarkan pada indikator moneter, non-moneter dan campuran (Arsyad, 1999). Adanya perbedaan capaian hasil pembangunan antar daerah menunjukkan belum tercapainya pemerataan ekonomi di suatu negara yang kemudian menyebabkan masalah ketimpangan ekonomi (Prawidya, 2016). Di Indonesia, pemerataan pembangunan ekonomi masih belum terwujud sebab konsentrasi pembangunan yang terpusat di wilayah Jawa. Berdasarkan kontribusi Produk Domestik Regional Bruto (PDRB) antar wilayah, Badan Pusat Statistik mencatat perekonomian Indonesia didominasi oleh kontribusi wilayah Jawa terhadap PDRB nasional sebesar 59,75%. Kemudian kontribusi wilayah Sumatera sebesar 21,36% diikuti oleh wilayah lainnya (BPS, 2021). Meskipun wilayah Sumatera berkontribusi cukup besar terhadap perekonomian Indonesia namun masih terdapat ketimpangan ekonomi antar daerah di Sumatera. Hal tersebut berdasarkan hasil penelitian Ridho (2022) yang menunjukkan terdapat beberapa provinsi dengan tingkat ketimpangan ekonomi tinggi dan sedang pada periode analisis tahun 2010-2019.

Pembangunan ekonomi berkaitan erat dengan pertumbuhan ekonomi yang merupakan salah satu indikator moneter keberhasilan pembangunan. Berdasarkan pendapatan per kapita, laju pertumbuhan ekonomi di wilayah Sumatera pada tahun 2020 mengalami perlambatan yang menunjukkan penurunan pendapatan daerah pada semua provinsi wilayah Sumatera. Hal tersebut berkaitan dengan pandemi COVID-19 yang menghambat pergerakan pembangunan ekonomi akibat pembatasan aktivitas masyarakat. Perlambatan laju ekonomi tersebut ditunjukkan seperti pada Gambar 1.



Gambar 1. Laju PDRB ADHK Menurut Provinsi di Sumatera (Sumber: BPS, 2021)

Selain itu, berdasarkan pada Peraturan Presiden Nomor 18 Tahun 2020 tentang Rancangan Pembangunan Menengah Jangka Panjang (RPJMN) tahun 2020-2024 di wilayah Sumatera masih terdapat banyak kabupaten/kota yang belum mencapai target pembangunan. Pada tahun 2024 wilayah Sumatera menargetkan laju pertumbuhan PDRB Sumatera 5,58% per tahun, IPM wilayah Sumatera mencapai kategori tinggi dengan nilai berkisar 71,90-78,19 tahun, persentase penduduk miskin wilayah Sumatera 5,76%, serta meningkatnya persentase rumah tangga yang mendapatkan akses air minum bersih dan layak menjadi 100%. Oleh sebab itu, perlu dilakukan peninjauan serta penentuan kebijakan khusus untuk daerah pembangunan ekonomi rendah dengan melakukan pengelompokan kabupaten/kota.

Penelitian ini bertujuan untuk mengetahui kelompok daerah dan membandingkan kinerja pengelompokan menggunakan analisis kluster. Analisis kluster yang digunakan dalam penelitian ini adalah metode *K-Medoids* dan *Clustering Large Application* (CLARA). Metode *K-Medoids* disebut juga dengan *Partitioning Around Method* (PAM) menjadikan objek amatan yang representatif (medoid) sebagai pusat kluster. CLARA merupakan metode kluster partisi berbasis sampling yang memanfaatkan algoritma *K-Medoids* pada beberapa data sampel sehingga waktu komputasi semakin efisien. Namun demikian metode CLARA memiliki kelemahan karena efektivitasnya bergantung pada ukuran data dan berbias pada penarikan sampel. Akan tetapi baik metode *K-Medoids* dan CLARA keduanya *robust* terhadap pencilan (Kaufman dan Rousseeuw, 2005; Saket, 2016).

Penelitian mengenai pengelompokan daerah berdasarkan indikator pembangunan ekonomi pernah dilakukan sebelumnya. Haryono (2016) dalam penelitiannya mengenai pengelompokan kabupaten/kota di Provinsi Jawa Timur menggunakan metode *C-Means* dan *Fuzzy C-Means*, hasil yang diperoleh menunjukkan bahwa metode *C-Means* merupakan metode terbaik. Namun, metode yang digunakan pada penelitian tersebut merupakan metode dalam analisis kluster yang tidak *robust* terhadap pencilan. Di sisi lain, mengenai permasalahan yang sama serta mengaitkannya dengan manfaat potensi daerah Renaldy (2018) menggunakan metode *Similarity Weight and Filter Method* (SWFM) dalam penelitiannya. Kemudian perbandingan metode *K-Medoids* dan CLARA telah dilakukan sebelumnya oleh AS (2019) dan Firda (2017). AS (2019) melakukan klusterisasi Kabupaten/Kota Provinsi Sulawesi Selatan dan Barat Berdasarkan Angka Partisipasi Pendidikan SMA/SMK/MA. Hasil penelitian tersebut menyimpulkan bahwa kedua metode sama baik digunakan untuk pengelompokan kabupaten/kota Provinsi Sulawesi Selatan dan Barat. Hal ini menyebabkan metode yang digunakan tidak memberikan hasil yang optimum akibat data penelitian yang berukuran kecil. Sedangkan Firda (2017) melakukan pengelompokan kabupaten/kota di Indonesia berdasarkan Indeks Pembangunan Manusia. Dalam penelitian tersebut diperoleh hasil bahwa metode *K-Medoids* dan CLARA merupakan alternatif metode yang dapat digunakan untuk mengelompokkan kabupaten/kota di Indonesia. Meskipun demikian, *K-Medoids* memberikan hasil yang lebih efektif karena kestabilan anggota kelompok dalam pengacakan. Sedangkan metode CLARA lebih efisien secara komputasi.

II. METODE PENELITIAN

A. *K-Medoids*

Metode *K-Medoids* merupakan salah satu algoritma pada metode partisi dalam analisis kluster, yang disebut juga dengan PAM (*Partitioning Around Method*) membentuk kelompok dengan menjadikan objek amatan yang berada terpusat ditengah kluster sebagai medoid (o_j) yang merepresentasikan kelompok (Schubert, 2019). Oleh karena itu,

metode *K-Medoids* memiliki kelebihan yaitu *robust* terhadap data yang memiliki pencilan. Selain itu, metode *K-Medoids* merupakan algoritma yang mudah dan sederhana untuk diterapkan (Saket dan Pandya, 2016).

Adapun algoritma kluster menurut (Kaufman dan Rousseeuw, 2005) dengan menggunakan metode *K-Medoids* adalah sebagai berikut.

1. Menentukan jumlah kluster/kelompok yang akan dibentuk sebanyak k .
2. Memilih secara acak objek amatan (o_j) yang akan digunakan sebagai medoid.
3. Menempatkan objek amatan yang bukan medoid dalam kelompok yang memiliki jarak terdekat terhadap suatu pusat kluster dengan menggunakan perhitungan jarak *Euclidean*.

$$d(o_j, o_i) = \sqrt{\sum_{i=1}^n (o_j - o_i)^2} \quad (1)$$

Dimana $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$; o_j adalah medoid kelompok ke- j dan o_i adalah objek amatan ke- i

4. Menghitung total simpangan (S) kemudian lakukan kembali langkah 2 – 4, iterasi berhenti bila nilai.

$$S = \text{Total Jarak Baru} - \text{Total Jarak Lama} \quad (2)$$

B. Clustering Large Application (CLARA)

Clustering Large Application (CLARA) merupakan metode kluster partisi berbasis sampling yang menerapkan algoritma PAM untuk mengelompokkan objek amatan pada data sampel yang ditarik secara acak. Penerapan algoritma PAM terus dilakukan pada data sampel hingga diperoleh solusi terbaik. Dengan demikian, waktu komputasi menjadi lebih efisien (Kaufman dan Rousseeuw, 2005). Metode CLARA efektif digunakan pada data berukuran besar serta *robust* terhadap adanya pencilan. Namun, efektivitas metode CLARA bergantung pada ukuran data dan bias pada penarikan sampel (Saket, 2016). Adapun algoritma kluster menurut (Kaufman dan Rousseeuw, 2005) menggunakan metode CLARA adalah sebagai berikut:

1. Menentukan jumlah kluster/kelompok yang akan dibentuk sebanyak k .
2. Menarik data sampel sebanyak lima gugus sampel (n')

$$n' = 40 + 2k \quad (3)$$

3. Menentukan medoid secara acak dan menerapkan metode PAM pada data sampel.
4. Menempatkan populasi dari semua objek non medoid pada kelompok berdasarkan jarak kedekatannya dengan menggunakan perhitungan jarak *Euclidean*.

$$d(o_j, o_i) = \sqrt{\sum_{i=1}^n (o_j - o_i)^2} \quad (4)$$

Dimana $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$; o_j adalah medoid kelompok ke- j dan o_i adalah objek amatan ke- i

5. Menghitung total simpangan (S) pada hasil klusterisasi populasi kemudian lakukan kembali langkah 3 – 5, iterasi berhenti bila nilai $S > 0$.

C. Silhouette Coefficient

Silhouette Coefficient merupakan metode yang digunakan untuk dan mengukur kualitas hasil klusterisasi serta dapat menentukan jumlah kelompok yang optimal. Metode ini menggabungkan metode *cohesion* dan *separation*. Ukuran *cohesion* adalah jarak objek ke- i terhadap objek pusat kluster (medoid) itu sendiri. Sedangkan ukuran *separation* ialah jarak objek ke- i terhadap objek pusat kluster (medoid) terdekat pada kluster yang berbeda (Daniel dan Chantal, 2015). Nilai *Silhouette Coefficient* berkisar antara -1 dan 1 . Nilai 1 dapat diartikan bahwa seluruh objek atau titik dimasukkan pada gerombol dengan tepat. Kemudian nilai 0 diartikan bahwa objek berada diantara dua kelompok yang sama sedangkan nilai -1 diartikan bahwa objek dimasukkan pada kelompok yang tidak seharusnya. Nilai *Silhouette Coefficient* tertinggi diantara jumlah k lainnya menunjukkan jumlah kelompok yang optimum dibentuk (Kaufman dan Rousseeuw, 2005).

Adapun hasil klusterisasi yang memiliki nilai rata-rata *Silhouette Coefficient* tertinggi menunjukkan hasil pengelompokkan yang paling baik (Kaufman dan Rousseeuw, 2005). Nilai *Silhouette Coefficient* pada suatu objek amatan ke- i dihitung dengan persamaan berikut.

$$SC(i) = \frac{1}{n_k} \sum_{x_{ij} \in C_k} \frac{b_i - a_i}{\max\{a_i, b_i\}} \quad (5)$$

dimana $i = 1, 2, \dots, n_k$; $k = 1, 2, \dots$; x_{ij} adalah objek amatan ke- i variabel ke- j ; C_k adalah kluster ke- k ; b_i adalah ukuran *cohesion* objek ke- i dan a_i adalah ukuran *separation* objek ke- i .

D. Jenis dan Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder yang bersumber dari *website* dan publikasi Statistik Kesejahteraan Masyarakat tahun 2020 oleh Badan Pusat Statistik. Adapun variabel yang digunakan pada penelitian ini berdasarkan indikator pembangunan ekonomi ditampilkan pada Tabel 1.

Tabel 1. Variabel Penelitian

Nama	Keterangan	Skala	Satuan
X_1	Produk Domestik Regional Bruto perkapita	Rasio	Ribu Rupiah
X_2	Indeks Pembangunan Manusia	Rasio	Persen
X_3	Angka Partisipasi Sekolah Berusia 7-12	Rasio	Persen
X_4	Angka Partisipasi Sekolah Berusia 13-15	Rasio	Persen
X_5	Angka Partisipasi Sekolah Berusia 16-18	Rasio	Persen
X_6	Tingkat Pengangguran Terbuka	Rasio	Persen
X_7	Tingkat Partisipasi Angkatan Kerja	Rasio	Persen
X_8	Persentase Penduduk Miskin	Rasio	Persen
X_9	Sumber Penerangan PLN	Rasio	Persen
X_{10}	Sumber Air Minum Bersih	Rasio	Persen
X_{11}	Sumber Air Minum Layak	Rasio	Persen

Sumber:(Renaldy, 2018 dan Haryono, 2016)

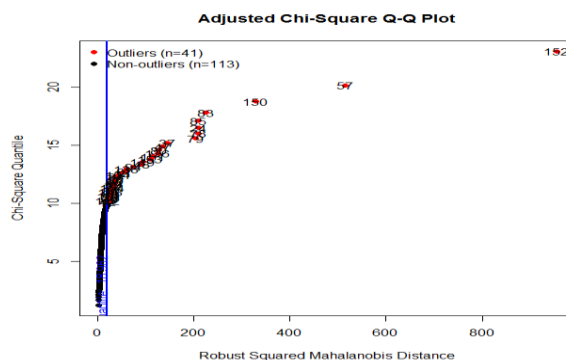
Adapun langkah-langkah analisis dalam penelitian adalah sebagai berikut:

1. Melakukan pemeriksaan mengenai keberadaan pencilan multivariat.
2. Melakukan standardisasi data dengan cara mengubah data kedalam bentuk standar baku.
3. Menentukan jumlah kluster/kelompok (k) optimal yang akan dibentuk menggunakan *Silhouette Coefficient* dengan Persamaan 4.
4. Melakukan pengelompokkan kabupaten/kota menggunakan metode *K-Medoids* dan CLARA dengan jarak pada Persamaan 1.
5. Memvalidasi hasil klasterisasi dengan membandingkan hasil kedua metode tersebut menggunakan metode *Silhouette Coefficient* pada Persamaan 4.
6. Menginterpretasi karakteristik dari hasil pengelompokkan terbaik.

III. HASIL DAN PEMBAHASAN

A. Mendeteksi Pencilan

Sebaran pencilan pada data pembangunan ekonomi wilayah Indonesia ditunjukkan seperti pada Gambar 2.



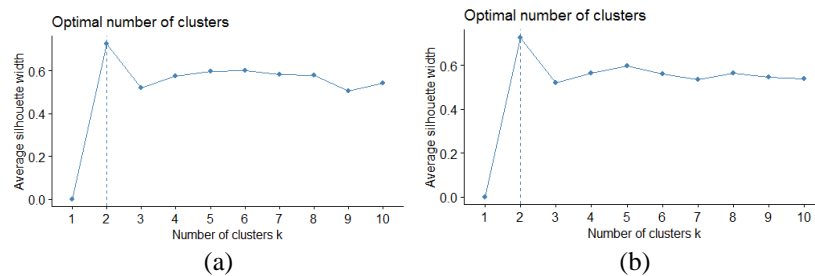
Gambar 2. Sebaran Pencilan Pembangunan Ekonomi

Berdasarkan pada Gambar 2 hasil perhitungan jarak Mahalanobis pada data pembangunan ekonomi diperoleh bahwa sebanyak 41 kabupaten/kota merupakan amatan pencilan yang terdiri dari 32 kabupaten dan 9 kota. Oleh

karena adanya pencilan pada data penelitian indikator pembangunan ekonomi maka analisis kluster yang tepat digunakan adalah analisis yang *robust* terhadap pencilan yaitu metode *K-Medoids* dan CLARA.

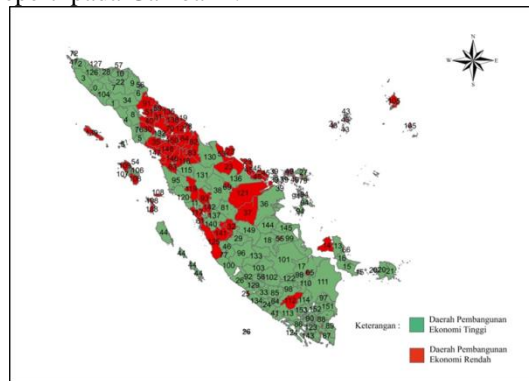
B. Analisis Kluster untuk mengelompokkan Kabupaten/Kota berdasarkan Indikator Pembangunan Ekonomi

Pada penelitian ini akan dibentuk kelompok yang optimal dengan nilai *Silhouette Coefficient* ditentukan berdasarkan nilai tertinggi. Jumlah kelompok yang optimal dibentuk berdasarkan nilai *Silhouette Coefficient* ditunjukkan seperti pada Gambar 3.



Gambar 3. (a) Plot *Silhouette Coefficient* *K-Medoids* dan (b). Plot *Silhouette Coefficient* CLARA

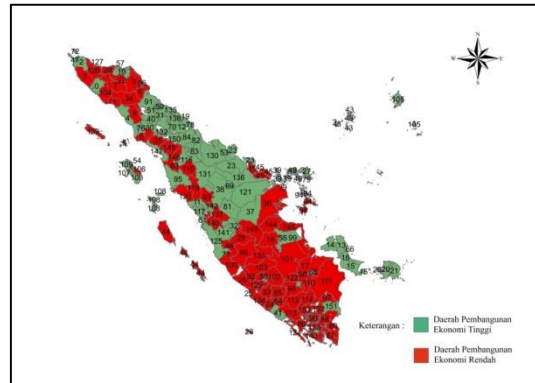
Berdasarkan Gambar 3 (a) dan (b) nilai *Silhouette* tertinggi ditunjukkan oleh jumlah kelompok 2 sehingga jumlah kelompok yang optimal dibentuk dengan metode *K-Medoids* dan CLARA adalah dua kelompok. Dengan demikian pengelompokan kabupaten/kota di wilayah Sumatera berdasarkan indikator pembangunan ekonomi dilakukan dengan membentuk dua kelompok pada metode *K-Medoids* dan CLARA. Selain itu jumlah kelompok yang optimal dibentuk dengan menggunakan nilai *Silhouette Coefficient* sebelumnya diperoleh sebanyak dua kelompok. Pengelompokan kabupaten/kota dengan metode *K-Medoids* dua kelompok diperoleh medoid yang terpilih ialah Kab. Aceh Barat dan Kab. Kepulauan Meranti. Metode ini *K-Medoids* memberikan hasil kelompok pertama sebanyak 59 kabupaten/kota dengan medoid Kab. Aceh Barat. Kemudian kelompok kedua sebanyak 95 kabupaten/kota dengan medoid Kab. Kepulauan Meranti. Adapun hasil pengelompokan kabupaten/kota dengan metode *K-Medoids* ditunjukkan seperti pada Gambar 4.



Gambar 4. Peta Sebaran Pengelompokan dengan Metode *K-Medoids* ($k = 2$)

Berdasarkan Gambar 4 daerah pada kelompok pertama ditunjukkan oleh daerah kabupaten/kota berwarna merah, yang terindikasi sebagai daerah dengan pembangunan ekonomi rendah. Daerah kabupaten/kota tersebut paling banyak terdapat di Provinsi Sumatera Utara, kemudian beberapa daerah kabupaten/kota di Provinsi Sumatera Barat dan Riau, serta provinsi lainnya. Adapun daerah pada kelompok kedua ditunjukkan oleh daerah kabupaten/kota berwarna hijau, yang merupakan daerah pembangunan ekonomi tinggi.

Pengelompokan kabupaten/kota dengan metode CLARA 2 kelompok diperoleh medoid yang terpilih ialah Kab. Labuhan dan Kab. Musi Rawas. Metode ini menghasilkan kelompok pertama sebanyak 74 kabupaten/kota dan kelompok kedua sebanyak 80 kabupaten/kota. Adapun hasil pengelompokan kabupaten/kota dengan metode CLARA ditunjukkan seperti pada Gambar 5.



Gambar 5. Peta Sebaran Pengelompokkan dengan Metode CLARA ($k = 2$)

Berdasarkan Gambar 5 daerah pada kelompok pertama ditunjukkan oleh daerah kabupaten/kota berwarna hijau, yang terindikasi sebagai daerah dengan pembangunan ekonomi tinggi sedangkan daerah pada kelompok kedua ditunjukkan oleh daerah kabupaten/kota berwarna merah, yang merupakan daerah pembangunan ekonomi rendah. Daerah pembangunan ekonomi rendah tersebar dibagian utara, barat dan selatan pulau Sumatera terutama kabupaten/kota di Provinsi Aceh dan sebagian besar daerah selatan Pulau Sumatera yaitu Provinsi Jambi, Bengkulu, Sumatera Selatan dan Lampung.

Hasil pengelompokkan terbaik ditunjukkan berdasarkan nilai *Silhouette Coefficient* tertinggi. Adapun nilai *Silhouette Coefficient* pengelompokkan kabupaten/kota pada penelitian ini ditunjukkan seperti pada Tabel 3.

Tabel 3. Nilai *Silhouette Coefficient*

Metode	SC_1	SC_2	SC
<i>K-Medoids</i> $k = 2$	0,06	0,18	0,13
CLARA $k = 2$	0,16	0,13	0,15

Berdasarkan Tabel 3 metode CLARA memberikan nilai *Silhouette Coefficient* sebesar 0,15 dan lebih tinggi dibandingkan dibandingkan metode *K-Medoids* sebesar 0,13. Namun, perbedaan penggerombolan pada kedua metode tidak jauh berbeda. Perbedaan nilai *Silhouette Coefficient* yang sangat kecil antara *K-Medoids* dan CLARA mengindikasikan kemiripan dari kedua metode tersebut relatif kecil. Dengan demikian, baik metode *K-Medoids* dan CLARA memberikan hasil mengelompokkan dengan kinerja yang sama. Hasil pengelompokkan metode CLARA memberikan hasil yang stabil dilihat pada nilai *Silhouette Coefficient* pada kelompok 1 sebesar 0,16 dan kelompok 2 sebesar 0,13. Oleh sebab itu, pada penelitian ini akan diinterpretasikan karakteristik dan hasil pengelompokkan dengan metode CLARA. Adapun karakteristik kelompok klaster metode CLARA ditunjukkan seperti pada Tabel 4.

Tabel 4. Karakteristik Kelompok Pembangunan Derah

Variabel	Rendah	Tinggi
PDRB per Kapita	38442,29 ²	64938,76 ¹
IPM	67,68 ²	73,46 ¹
Tingkat pengangguran terbuka	4,07 ¹	6,66 ²
Tingkat partisipasi angkatan kerja	71,51 ¹	66,88 ²
Persentase penduduk miskin	13,27 ²	8,18 ¹
Angka partisipasi sekolah umur 7 – 12 tahun	99,60 ¹	99,59 ²
Angka partisipasi sekolah umur 13 – 15 tahun	95,95 ²	96,69 ¹
Angka partisipasi sekolah umur 16 – 18 tahun	75,96 ²	79,85 ¹
Persentase sumber penerangan listrik PLN	94,62 ²	97,22 ¹
Persentase sumber air minum bersih	55,64 ²	84,37 ¹

Ket : ^{1,2} menunjukkan urutan nilai seharusnya

Berdasarkan Tabel 4 hasil pengelompokan kabupaten/kota dengan metode CLARA memiliki kategori rendah pada kelompok pertama dan kategori tinggi pada kelompok kedua. Hal tersebut dilakukan berdasarkan pada perbedaan nilai karakteristik. Pada kelompok 1 nilai variabel cenderung lebih rendah dibandingkan kelompok 2 kecuali variabel TPT, TPAK, dan APS 7 – 12. Hal ini menunjukkan kabupaten/kota yang berada pada kelompok 1 memiliki pembangunan ekonomi yang relatif rendah. Sebaliknya daerah yang berada pada kelompok kedua memiliki pembangunan ekonomi yang cenderung tinggi dan baik. Berdasarkan peta sebaran pengelompokan dengan metode CLARA pada Gambar 5 terlihat bahwa daerah dengan pembangunan ekonomi tinggi cenderung terdapat pada wilayah timur Pulau Sumatera.

IV. KESIMPULAN

Hasil pengelompokan kabupaten/kota berdasarkan pembangunan ekonomi di wilayah terdiri dari 2 kategori. Kelompok pertama memiliki kategori daerah dengan pembangunan ekonomi rendah pada 74 kabupaten/kota di wilayah Sumatera. Sedangkan kelompok kedua memiliki kategori daerah dengan pembangunan ekonomi tinggi pada 80 kabupaten/kota di wilayah Sumatera. Hal tersebut menunjukkan proporsi daerah yang dengan pembangunan ekonomi rendah hampir sama banyak dengan daerah yang berlabel pembangunan ekonomi tinggi. Kelompok daerah pembangunan ekonomi rendah seharusnya menjadi perhatian khusus bagi pemerintah. Daerah kabupaten/kota tersebut tersebar dibagian utara, barat dan selatan pulau Sumatera terutama kabupaten/kota di Provinsi Aceh, Jambi, Bengkulu, Sumatera Selatan dan Lampung. Adapun saran kepada penelitian dengan permasalahan data pencilan untuk menggunakan metode pengelompokan yang lain seperti CLARANS, CURE dan sebagainya.

DAFTAR PUSTAKA

- Arsyad, L. (2010). *Ekonomi Pembangunan*. Yogyakarta: UPP STIM YKPN
- AS, W., Aidid, M. K., dan Nusrang, M. (2019). Pengelompokan Kabupaten/Kota Provinsi Sulawesi Selatan dan Barat Berdasarkan Angka Partisipasi Pendidikan SMA/SMK/MA Menggunakan *K-Medoids* dan CLARA. *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, 1(3), 48-65. DOI:10.35580/variasiunm12899
- Badan Pusat Statistik. (2021). *Laporan Perekonomian Indonesia 2021*, Badan Pusat Statistik, Jakarta: BPS.
- Daniel, L. T., dan Chantal, L. T. (2015). *Data Mining ang Predictive Analytics*. New Jersey: John Wiley dan Sons, Inc.
- Firda, R. A. (2018). Penggerombolan Daerah di Indonesia dengan K-Medoid dan CLARA (Studi Kasus Peubah Indeks Pembangunan Manusia Tahun 2015), Skripsi, Institut Pertanian Bogor, Bogor, Indonesia, Oktober 2018
- Haryono, Cendiana Aprilia. (2016). Penerapan metode C-Means dan Fuzzy C-Means pada Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Pembangunan Ekonomi, Skripsi, Institut Teknologi Sepuluh Nopember.
- Kaufman, L., dan Rousseeuw, P. J. (2005). *Finding Groups in Data An Introduction to Cluster Analysis*. New Jersey: John Wiley dan Sons, Inc.
- Peraturan Presiden Nomor 18 Tahun 2020 tentang Rencana Pembangunan Jangka Menengah Nasional (RPJMN) Tahun 2020-2024
- Prawidya. (2016). Analisis Ketimpangan Ekonomi dan Pengaruhnya terhadap Tingkat Kriminalitas di Propinsi Sumatera Utara. *Jurnal Ekonomikawan*, 16.
- Renaldy, A L., dan Zain, I., (2018). "Pengelompokan Kabupaten/Kota Berdasarkan Indikator Pembangunan Ekonomi dan Potensi Daerah Provinsi Jawa Timur Menggunakan *Similarity Weight and Filter Method (SWFM)*", *INFERENSI*, Vol. 1, No. 2, hal. 49-56.
- Ridho, M., dkk. (2022). Analisis ketimpangan pembangunan wilayah antar provinsi di Pulau Sumatera. *E-Jurnal Perspektif Ekonomi Dan Pembangunan Daerah*, 11(1), 47-58. <https://doi.org/10.22437/pdpd.v11i1.14005>
- Saket, S., dan Pandya, S. (2016). An Overview of Partitioning Algorithms in Clustering Techniques. *International Journal of Advanced Research in Computer Engineering dan Technology (IJARCET)*, 5(6).