Pengelompokkan Desa di Jawa Barat Berdasarkan Variabel Penyusun Indeks Desa Membangun

Via Fauziah*, Suliadi

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*viafauziah26@gmail.com, suliadi@gmail.com

Abstract. The development and improvement of the quality of life at the village level is a crucial aspect of a country's development. The Village Development Index (IDM) is one of the important indicators that describe the development of villages from various aspects. The IDM measurement contains three constituent indices, namely the Social Resilience Index (IKS), Economic Resilience Index (IKE), and Environmental Resilience Index (IKL). IDM data in 2023 for West Java Province have significant variations in the values of the three constituent indices. Therefore, a clustering method is required to determine groups consisting of several villages in West Java. The clustering method that is often used is the K-Means method, but K-Means has a weakness, which is sensitive to outliers. Therefore, K-Medoids is present to overcome these shortcomings. The results of research with the Silhouette method obtained the optimal number of clusters is as many as 2. So that clustering is carried out with a number of clusters of 2. The results of clustering with the K-Medoids method were: (1) Cluster 1 had an average values of IKS, IKE, and IKL variables that were below the population average. So that cluster 1 had a lower IDM value category compared to cluster 2. (2) Cluster 2 had an average values of the IKS, IKE, and IKL variables that are above the population average. Therefore, cluster 2 has a higher IDM category than cluster 1.

Keywords: *Village Development Index, Clustering, K-Means, K-Medoids.*

Abstrak. Pembangunan dan peningkatan kualitas kehidupan di tingkat desa merupakan aspek krusial dalam pembangunan suatu negara. Indeks Desa Membangun (IDM) menjadi salah satu indikator penting yang menggambarkan perkembangan desa dari berbagai aspek. Pengukuran IDM memuat tiga indeks penyusun, yakni Indeks Ketahanan Sosial (IKS), Indeks Ketahanan Ekonomi (IKE), dan Indeks Ketahanan Lingkungan (IKL). Berdasarkan data IDM tahun 2023, terlihat bahwa desa-desa di Provinsi Jawa Barat memiliki variasi nilai yang cukup signifikan dalam ketiga indeks penyusunnya. Untuk itu, diperlukan metode klasterisasi untuk menentukan kelompok yang terdiri dari beberapa desa di Jawa Barat. Metode klastering yang sering digunakan adalah metode *K-Means* Namun *K-Means* memiliki kelemahan, yaitu sensitif terhadap *outlier*. Maka dari itu, *K-Medoids* hadir untuk mengatasi kekurangan tersebut. Adapun hasil penelitian dengan metode Silhouette diperoleh jumlah klaster optimal adalah sebanyak 2. Sehingga klasterisasi dilakukan dengan jumlah klaster sebanyak 2. Hasil klasterisasi dengan metode *K-Medoids* adalah : (1) Klaster 1 memiliki nilai rata-rata untuk variabel IKS, IKE, dan IKL yang berada di bawah rata-rata populasi. Sehingga klaster 1 memiliki kategori nilai IDM yang lebih rendah dibandingkan dengan klaster 2. (2) Klaster 2 memiliki nilai rata-rata untuk variabel IKS, IKE, dan IKL yang berada di atas rata-rata populasi. Sehingga klaster 2 memiliki kategori IDM yang lebih tinggi dibandingkan dengan klaster 1.

Kata Kunci: *Indeks Desa Membangun, Klasterisasi, K-Means, K-Medoids.*

Corresponding Author
Email: suliadi@gmail.com

A. Pendahuluan

Menurut Undang-Undang Nomor 6 Tahun 2014 tentang desa, desa diartikan sebagai kesatuan masyarakat hukum yang mempunyai batas wilayah yang bertugas untuk mengatur dan mengurus urusan pemerintahan, kepentingan masyarakat setempat sesuai prakarsa masyarakat, hak asal usul, dan atau hak tradisional yang diakui dan dihormati dalam sistem pemerintahan Negara Kesatuan Republik Indonesia (Sahwa Chanigia Viqri & Kurniati, 2023).

Indeks Desa Membangun (IDM) menjadi salah satu indikator penting yang menggambarkan perkembangan desa dari berbagai aspek. Pembangunan yang merata di seluruh wilayah desa memerlukan pemahaman yang mendalam terhadap kondisi setiap desa. Indeks Desa Membangun memotret perkembangan kemandirian desa berdasarkan implementasi Undang-Undang Desa dengan dukungan Dana Desa serta Pendamping Desa. [1]

Jawa Barat adalah salah satu Provinsi di Indonesia yang memiliki tingkat pencapaian yang baik dalam nilai Indeks Desa Membangun. Melalui data IDM tahun 2023, terlihat bahwa desa-desa di Provinsi Jawa Barat memiliki variasi nilai yang cukup signifikan dalam ketiga indeks penyusunnya. Ragam nilai indeks ini menunjukkan adanya pola yang mungkin dapat memberikan informasi penting kepada pemerintah, baik sebagai evaluasi maupun landasan untuk kebijakan pembangunan desa yang lebih terperinci (Khofifah, 2022).

Klasterisasi memungkinkan pembentukan sejumlah kelompok atau klaster yang terdiri dari beberapa desa di Jawa Barat. Salah satu metode klasterisasi yang dapat digunakan adalah metode K-Means. Alasan penggunaan metode K-means diantaranya karena metode ini memiliki ketelitian yang cukup tinggi terhadap ukuran objek, sehingga metode ini relatif lebih terukur dan efisien untuk pengolahan objek dalam jumlah besar [2]. Namun K-Means memiliki kelemahan, yaitu sensitif terhadap outlier. Maka dari itu K-Means kurang cocok digunakan pada data yang memiliki outlier (Oktoriandi, 2022).

K-Medoids merupakan modifikasi dari K-Means, dimana K-Medoids hadir untuk mengatasi kekurangan pada K-Means yang sensitif terhadap pencilan karena menggunakan nilai rata-rata (mean) sebagai pusat klasternya [3]. K-Means menggunakan centroid sebagai pusat klaster, sedangkan K-Medoids menggunakan medoid sebagai pusat klaster. Maka dari itu apabila data memiliki outlier, K-Medoids mampu melakukan klasterisasi dengan lebih baik.

Berdasarkan latar belakang yang telah diuraikan, maka tujuan dalam penelitian ini adalah :

1. Menentukan jumlah klaster optimal pada klasterisasi desa di Jawa Barat berdasarkan data Indeks Desa Membangun tahun 2023
2. Melakukan klasterisasi desa di Jawa Barat dengan menggunakan metode K-Medoids berdasarkan data Indeks Desa Membangun tahun 2023

B. Metodologi Penelitian

Peneliti menggunakan metode klasterisasi *K-Medoids*, di mana data yang digunakan merupakan data sekunder dari tahun 2023 tentang Indeks Desa Membangun (IDM) Jawa Barat tahun 2023. Data diperoleh dari website https://portaldatadesa.jabarprov.go.id/ milik Dinas Pemberdayaan Masyarakat dan Desa Provinsi Jawa Barat. Data Indeks Desa Membangun ini dikelola oleh KDPDTT (Kementerian Desa, Pembanguna Daerah Tertinggal, dan Transmigrasi Republik Indonesia). Objek yang akan diteliti yaitu desa di Jawa Barat tahun 2023 sebanyak 5311 desa. Variabel bebas dalam penelitian ini adalah variabel penyusun IDM tahun 2023, yaitu: (1) nilai Indeks Ketahanan Sosial (IKS); (2) nilai Indeks Ketahanan Ekonomi (IKE); (3) nilai Indeks Ketahanan Lingkungan/ ekologi (IKL).

Klasterisasi

Klasterisasi adalah suatu teknik atau metode untuk mengelompokkan data. Klasterisasi adalah sebuah proses untuk mengelompokan pengamatan ke dalam beberapa klaster atau kelompok sehingga data dalam satu klaster memiliki tingkat kesamaan yang maksimum dan data antar klaster memiliki kesamaan yang minimum [2]. Klasterisasi merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan klaster. Objek yang ada di dalam klaster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan klaster

yang lain. Terdapat dua jenis metode analisis klasterisasi yang digunakan dalam pengelompokan data, yaitu klaster hierarki dan klaster non hirarki. Berikut penjelasan langkah-langkah metode analisis klasterisasi [4].

1. Klaster hierarki adalah metode analisis pengelompokan data yang dimulai dengan mengelompokkan 2 atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan berikutnya. Demikian seterusnya sehingga klaster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah klaster/kelompok. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut. Contoh metode dari klaster hierarki adalah *Linkage Method, Ward's Method, Automatic Interaction Detection* (AID) dan lain sebagainya.
2. Klaster non hierarki adalah metode pengelompokan data yang dimulai dengan menentukan terlebih dahulu jumlah klaster/kelompok yang diharapkan. Kemudian, dilanjutkan proses klaster/kelompok tanpa mengikuti proses hierarki. Contoh metode dari klaster non hierarki adalah *K-Means, K-Medoids, Gaussian Mixture Model* dan lain sebagainya.

K-Medoids

K-medoids atau dikenal juga dengan metode *Partitioning Around Medoids* (PAM), dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw. *K-medoids* adalah algoritma pengelompokan yang memiliki kemiripan dengan *K-means*, namun tidak terlalu sensitif terhadap outlier/pencilan [5]. *K-medoids* adalah sebuah teknik terkenal dalam pengelompokan non-hierarki yang merupakan varian dari metode *K-Means* [6]. *K-Medoids* memiliki metode yang lebih baik karena lebih kuat terhadap *noise* dan *outlier* [7].

K-means menggunakan *centroid* sebagai pusat klaster. Sedangkan pada algoritma *K-medoid* menggunakan medoid sebagai pusat klaster [8]. *K-Means* sangat sensitif terhadap data yang mengandung pencilan. Sebaliknya, algoritma *K-Medoids* kurang sensitif terhadap pencilan karena menggunakan medoid, yang membuat prosedur komputasi pada *K-Medoids* lebih kompleks daripada prosedur pada algoritma *K-Means* [9]. Alih-alih menggunakan nilai rata-rata objek di setiap klaster, algoritma *K-Medoids* menggunakan objek pada kumpulan objek sebagai titik pusat [5]. Secara sederhana, definisi dari suatu medoid adalah amatan antara anggota klaster dengan rata-rata jarak ke amatan lain yang paling kecil [10].

Medoids dari klaster dapat didefinisikan sebagai berikut [10] :

$$medoids_l = x_k \text{ dengan } k = argmin\ d_i \quad (1)$$

x_k tidak lain adalah pengamatan ke-k. Sehingga medoids untuk kluster l atau $medoids_l$, tidak lain adalah pengamatan ke-k, dimana pengamatan ke-k adalah pengamatan yang memiliki rata-rata jarak paling minimum terhadap obyek lain pada klaster ke-l.

Berikut langkah-langkah melakukan pengelompokkan dengan menggunakan metode *K-Medoids* :

1. Menentukan jumlah klaster k yang ingin dibentuk
2. Inisialisasi pusat klaster (medoids awal) sebanyak k dari n data.
3. Hitung jarak setiap obyek ke masing-masing medoid dan alokasikan setiap objek ke dalam klaster berdasarkan jarak terkecil ke medoid
4. Hitung total jarak terdekat ke medoid
5. Pilih secara acak objek pada masing-masing klaster sebagai kandidat medoid baru (non medoid).
6. Menghitung jarak setiap objek yang berada pada masing-masing klaster dengan medoid baru.
7. Hitung total simpangan (SM) dengan menghitung:
   $SM = b - a \quad (2)$
   dimana b merupakan nilai total jarak baru dan a merupakan total jarak lama. Jika $SM <$ 0, maka tukar objek dengan data klaster lainnya untuk membentuk sekumpulan k objek baru sebagai medoid.

8. Ulangi langkah 5 sampai 7 hingga tidak terjadi perubahan medoid, sehingga didapat klaster beserta anggotanya masing-masing

Sillhoutte Index

Salah satu metode untuk menentukan jumlah klaster optimal adalah dengan metode *Silhouette* [10]. *Sillhoutte Index* adalah ukuran seberapa mirip sebuah objek dengan klasternya sendiri dibandingkan dengan klaster lain. Indeks ini dilihat dengan mengukur derajat kepercayaan dalam proses pengklasterannya. Nilai 1 menunjukkan yang terbaik, yang berarti bahwa titik data i sangat kompak di dalam klaster tempat ia berada dan jauh dari klaster lainnya.

Misalkan klaster ke-l beranggotakan n_l objek/pengamatan, dan i adalah objek yang ada pada klaster c_l untuk i = 1,2,…n_l. Rata-rata jarak objek ke-i terhadap objek lainnya pada klaster yang sama dapat dihitung sebagai berikut :

$$a(i) = \frac{1}{|n_l|-1}\sum_{i\neq i'} d(i,i') \tag{3}$$

dimana $d(i,i')$ adalah jarak objek ke-i terhadap objek i' pada klaster yang sama. $a(i)$ menggambarkan seberapa dekat objek ke-i terhadap objek lainnya pada klaster yang sama.

Rata-rata ketidakmiripan (disimiliarity) objek ke-i pada klaster c_l terhadap klaster lain yaitu l' didefinisikan sebagai $b(i)$ yang merupakan rata-rata jarak i ke semua pengamatan/objek yang ada di klaster l'. Untuk nilai b_i dapat dihitung sebagai berikut :

$$b(i) = \min\left\{\frac{1}{|nl'|}\sum_{r\in cl'} d(i,r)\right\}_{l\neq l';l'=1,2,\dots,k} \tag{4}$$

dimana d(i,r) adalah jarak objek ke-i klaster ke-l dengan objek ke-r klater ke-l'.

Jika $a(i)$ semakin kecil, maka semakin bagus karena itu berarti objek sudah tempatkan dengan baik pada klaster tersebut. Sedangkan jika $b(i)$ semakin besar, maka semakin bagus. Adapun perhitungan *Sillhoutte Index* untuk objek ke-i adalah sebagai berikut :

$$s(i) = \frac{b(i)-a(i)}{\max((a(i),b(i))} \tag{5}$$

dimana $a(i)$ adalah rata-rata jarak objek ke-i terhadap objek lainnya pada klaster yang sama. Kisaran nilai $s(i)$ adalah $-1 \leq s(i) \leq 1$. Dengan demikian, $s(i)$ yang mendekati 1 berarti bahwa data dikelompokkan dengan tepat. Jika mendekati -1, maka dengan logika yang sama kita melihat bahwa i akan lebih cocok jika dikelompokkan dalam klaster tetangganya. Nilai $s(i)$ yang mendekati nol berarti data tersebut berada di perbatasan klaster.

Sillhoutte Index atau nilai silhouette untuk banyaknya klaster k, dinotasikan dengan $\tilde{S}(k)$, yang didefinisikan sebagai rata-rata $s(i)$ untuk seluruh data $(i = 1,2,\dots,n)$ jika obyek diklasterkan menjadi k klaster. Nilai ini mengukur seberapa tepat data tersebut diklasterkan. Nilai *silhoutte* $\tilde{S}(k)$ diperoleh melalui rumus :

$$\tilde{S}(k) = \frac{\sum_{i=1}^{n} s(i)}{n}. \tag{6}$$

Karena $-1 \leq s(i) \leq 1$, maka $-1 \leq \tilde{S}(k) \leq 1$. Semakin dekat nilai $\tilde{S}(k)$ dengan 1, menunjukkan semakin baiknya klaster tersebut. Sehingga plot *Silhoutte* dapat digunakan untuk menentukan banyaknya klaster yang tepat, yaitu nilai $\tilde{S}(k)$ yang paling besar yang paling dekat dengan 1.

C. Hasil Penelitian dan Pembahasan

Deskripsi Data

Pada bagian ini dilakukan analisis deskriptif data untuk variabel-variabel yang digunakan pada penelian ini.

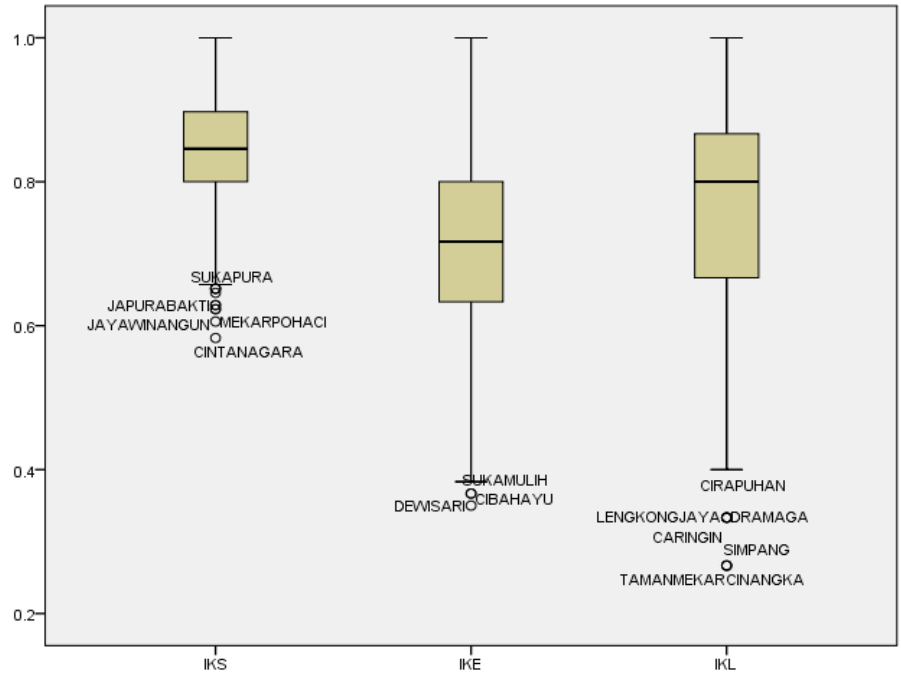
Tabel 1. Statistika Deskriptif

Variabel	Minimum	Maksimum	Rata-Rata	Std. Dev.	Koef. Keragaman
IKS	0,5829	1,0000	0,846262	0,071	8,36%
IKE	0,3500	1,0000	0,724499	0,114	15,74%
IKL	0,2667	1,0000	0,768700	0,134	17,45%

Berdasarkan Tabel 1, diperoleh nilai minimum, maksimum, rata-rata dan standar deviasi dari ketiga variabel penyusun Indeks Desa Membangun tahun 2023. Pada variabel IKS memperoleh nilai minimum 0,5829, nilai maksimum 1,0000, rata-rata 0,846262 dan standar deviasi 0,071. Pada variabel IKE memperoleh nilai minimum 0,3500, nilai maksimum 1,0000, nilai rata-rata 0,724499 dan standar deviasi 0,114. Pada variabel IKL memperoleh nilai minimum 0,2667, nilai maksimum 1,0000, rata-rata 0,768700 dan standar deviasi 0,134. Selain itu tampak bahwa IKE merupakan indeks dengan rata-rata yang paling rendah di antara ketiga indeks, disusul IKL kemudian IKS. Dilihat dari variabilitas, IKS memiliki variabilitas paling kecil dengan koefisien keragaman 8,36% yang menunjukkan IKS antar desa di Indonesia tidak terlalu berbeda. Sedangkan IKL memiliki keragaman yang cukup besar dengan koefisien keragaman sebesar 17,45%. Ini menunjukkan bahwa IKL desa-desa di Indonesia lebih bervariasi dibandingkan IKS dan IKL.

Identifikasi Outlier

Dalam penelitian ini, dilakukan identifikasi outlier untuk ketiga variabel yaitu IKS, IKE, dan IKL dengan menggunakan boxplot yang disajikan pada Gambar 1. Berdasarkan Gambar 1, terdapat pencilan di ketiga variabel. Pada Variabel IKS, terdapat pencilan pada data ke-4627 atau Sukapura dengan nilai 0,6229, data ke-5004 atau Japurabakti dengan nilai 0,6286, data ke-3840 atau Jayawinangun dengan nilai 0,6229, data ke-5263 atau Mekarpohaci dengan nilai 0,6057, dan data ke-2606 atau Cintanagara dengan nilai 0,5829. Untuk variabel IKE, terdapat pencilan pada data ke-4424 dengan nilai 0,3667, data ke-4930 atau Cibahayu dengan nilai 0,3667, dan data ke-5153 atau Dewisari dengan nilai 0,3500. Untuk variabel IKL, terdapat pencilan pada data ke-4795 atau Cirapuhan dengan nilai 0,3333, data ke-4905 atau Lengkongjaya dengan nilai 0,3333, data ke-4423 atau Dramaga dengan nilai 0,3333, data ke-4503 atau Caringin dengan nilai 0,3333, data ke-4728 atau Simpang dengan nilai 0,2667, data ke-5146 atau Taman Mekar dengan nilai 0,2667, dan data ke-4389 atau Cinangka dengan nilai 0,2667.



Gambar 1. Boxplot untuk Ketiga Variabel

Pemeriksaan Multikolinearitas

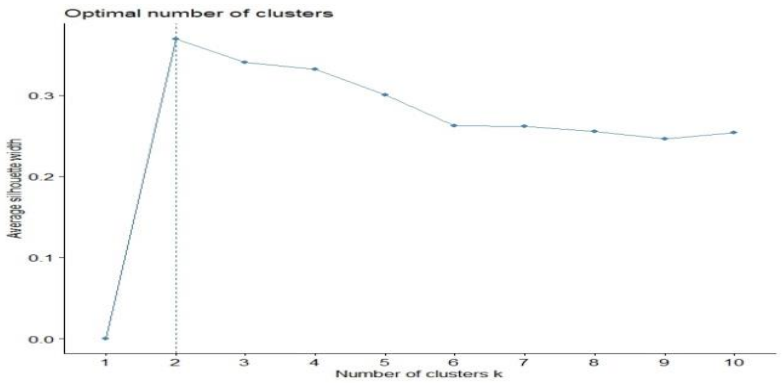
Untuk mendeteksi ada atau tidaknya masalah multikolinearitas bisa menggunakan *Variance Inflation Factors* (VIF). Apabila nilai $VIF > 10$ maka mengindikasikan ada masalah

multikolinieritas. Hasil perhitungan nilai VIF disajikan pada Tabel 2.

Tabel 2. Pemeriksaan Multikolinearitas

Variabel	VIF	Keterangan
IKS	1,377	Tidak ada masalah Multikolinearitas
IKE	1,027	Tidak ada masalah Multikolinearitas
IKL	1,007	Tidak ada masalah Multikolinearitas

Berdasarkan Tabel 2, masing-masing variabel memiliki nilai VIF < 10 sehingga dapat disimpulkan bahwa tidak terjadi multikolineritas.



Gambar 2. Pemilihan Jumlah Klaster Optimal

Pemilihan Jumlah Klaster Optimal

Pemilihan klaster optimal dilakukan dengan metode *Silhouette*. Jumlah klaster optimal dapat dilihat pada Gambar 2. Berdasarkan Gambar 2, dapat dilihat bahwa nilai k yang optimal adalah k = 2, karena memiliki rata-rata nilai *Silhoutte* tertinggi. Maka, dapat disimpulkan jumlah klaster yang optimal menggunakan metode *Silhouette* sebanyak 2 klaster. Jadi, proses klastering dilakukan dengan jumlah klaster sebanyak 2.

Hasil Klastering dengan K-Medoids

Tabel 3 berikut ini menyajikan medoid pada klaster 1 dan klaster 2.

Tabel 3. Medoid Untuk Setiap Klaster

Klaster	ID	IKS	IKE	IKL
1	2357	0.8629	0.7167	0.8667
2	1698	0.9086	0.8000	0.8000

Berdasarkan Tabel 3, medoid untuk klaster 1 terletak pada data ke-2357 atau desa Saganten dengan nilai medoid (0,8629;0,7167;0,8667). Sedangkan medoid untuk klaster 2 terletak pada data ke-1698 atau desa Cipeundeuy dengan nilai medoid (0,9086;0,8000;0,8000). Selanjutnya jumlah anggota dari masing-masing klaster dapat dilihat pada Tabel 4 sebagai berikut :

Tabel **4.** Jumlah Anggota Masing-Masing Klaster

Klaster	Jumlah
1	2722
2	2589

Berdasarkan Tabel 4, dapat dilihat bahwa jumlah anggota pada klaster 1 sebanyak 2722 desa. Sedangkan jumlah anggota dari klaster 2 adalah sebanyak 2589 desa. Hasil pengelompokkan desa secara lebih rinci disajikan pada Tabel 5.

Tabel 5. Hasil Pengelompokkan Desa

No	Kabupaten	Kecamatan	Desa	Klaster
1	Kab. Bogor	Gunung Putri	Wanaherang	1
2	Kab. Bogor	Gunung Putri	Bojong Kulur	2
3	Kab. Bogor	Gunung Putri	Ciangsana	2
4	Kab. Bogor	Gunung Putri	Gunung Putri	2
5	Kab. Bogor	Gunung Putri	Bojong Nangka	2
6	Kab. Bogor	Gunung Putri	Tlajung Udik	1
7	Kab. Bogor	Gunung Putri	Cicadas	2
8	Kab. Bogor	Gunung Putri	Cikeas Udik	2
9	Kab. Bogor	Gunung Putri	Nagrak	1
10	Kab. Bogor	Gunung Putri	Karanggan	2
11	Kab. Bogor	Citeureup	Puspasari	1
⋮				⋮
5311	Kab. Bandung Barat	Gununghalu	Sukasari	1

Selanjutnya tabulasi klasifikasi status desa berdasarkan hasil klasterisasi dapat dilihat pada Tabel 6 sebagai berikut.

Tabel 6. Klasifikasi Status Desa Berdasarkan Hasil Klasterisasi

Klaster	Status			Total
	Berkembang	Maju	Mandiri	
1	551	1290	881	2722
2	379	1263	947	2589
Total	930	2553	1828	5311

Berdasarkan Tabel 6, pada klaster 1 terdapat 551 desa dengan status "Berkembang", 1290 desa dengan status "Maju", dan 881 desa dengan status "Mandiri". Sedangkan pada klaster 2 terdapat 379 desa dengan status "Berkembang", 1263 desa dengan status "Maju", dan 947 desa dengan status "Mandiri". Sehingga klaster 1 memiliki lebih banyak desa dengan status "Berkembang" dan status "Maju", sedangkan pada klaster 2 memiliki lebih banyak desa dengan status "Mandiri". Selanjutnya dilakukan profilisasi dengan mencari nilai rata-rata dari setiap variabel. Berikut nilai rata-rata variabel pada setiap klaster yang disajikan pada Tabel 7.

Tabel 7. Nilai Rata-Rata Variabel Pada Setiap Klaster

	Klaster 1	Klaster 2	Rata-rata Populasi
IKS	0,8446	0,8480	0,8463
IKE	0,7179	0,7314	0,7245
IKL	0,7614	0,7763	0,7687

Berdasarkan Tabel 7, masing-masing klaster dapat diinterpretasikan sebagai berikut:

1. Klaster 1
   Klaster 1 memiliki nilai rata-rata IKS, IKL dan IKE lebih rendah daripada rata-rata populasi. Sehingga klaster 1 memiliki desa-desa dengan karakteristik IKS, IKE, dan IKL lebih rendah daripada klaster 2. Maka dapat disimpulkan desa-desa di klaster 1 memiliki kategori nilai Indeks Desa Membangun yang lebih rendah dibandingkan dengan klaster 2.
2. Klaster 2
   Klaster 2 memiliki nilai rata-rata IKS, IKL dan IKE lebih tinggi daripada rata-rata populasi. Sehingga klaster 2 memiliki desa-desa dengan karakteristik IKS, IKE, dan IKL lebih tinggi daripada klaster 1. Maka dapat disimpulkan desa-desa pada klaster 2 memiliki kategori nilai Indeks Desa Membangun yang lebih tinggi dibandingkan dengan klaster 1.

D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Jumlah klaster yang optimal menggunakan metode *Silhouette* adalah sebanyak 2 klaster.
2. Hasil klasterisasi dengan metode K-Medoids untuk k=2 adalah klaster 1 terdiri dari 2722 desa, sedangkan klaster 2 terdiri dari 2589 desa. Klaster 1 memiliki nilai rata-rata untuk variabel IKS, IKE, dan IKL yang berada di bawah rata-rata populasi. Sehingga klaster 1 memiliki kategori Indeks Desa Membangun yang lebih rendah daripada klaster 2. Klaster 2 memiliki nilai rata-rata untuk variabel IKS, IKE, dan IKL yang berada di atas rata-rata populasi. Sehingga klaster 2 memiliki kategori Indeks Desa Membangun yang lebih tinggi daripada klaster 1.

Acknowledge

Penulis mengucapkan terima kasih kepada semua pihak yang telah membantu dan mendukung penulis dalam menyelesaikan penelitian sederhana ini.

Daftar Pustaka

[1] KDPDTT. (2023). *IDM* (Online). (https://idm.kemendesa.go.id/view/detil/6/faq, diakses 6 Januari 2023).

[2] Tan, P.N., Steinbach, M., Kumar, V. (2006) Introduction to Data Mining. Boston : Pearson Education.

[3] Han, J., Kamber, M., dan Pei, J. (2012). Data mining concepts and techniques third edition. University of Illinois at Urbana-Champaign Micheline Kamber Jian Pei Simon Fraser University..

[4] Santoso, S. (2010). Statistik Multivariat. Jakarta: Elex Media Komputindo

[5] Mohemad, R., Muhait, N. N. M., Noor, N. M. M., & Othman, Z. A. (2022). Performance analysis in text clustering using k-means and k-medoids algorithms for Malay crime documents. International Journal of Electrical and Computer Engineering (IJECE), 12(5), 5014-5026.

[6] Park, H. S., & Jun, C. H. (2009). A simple and fast algorithm for K-medoids clustering. *Expert systems with applications*, *36*(2), 3336-3341.

[7] Soni, K. G., & Patel, A. (2017). Comparative Analysis of K-means and K-medoids Algorithm on IRIS Data. International Journal of Computational Intelligence Research, 13(5), 899-906.

[8] Fitriyani, E. N., & Achmad, A. I. (2023, July). Penerapan Analisis K-Medoids Cluster untuk Mengelompokkan Wilayah di Provinsi Jawa Barat Berdasarkan Fasilitas Kesehatan Tahun 2021. In *Bandung Conference Series: Statistics* (Vol. 3, No. 2, pp. 283-293).

[9] Mingyung, L., Seoghun, L., Jaehwa, P., & Seongwon, S. (2020). Clustering and Characterization of the Lactation Curves of Dairy Cows Using K-Medoids Clustering Algorithm. Animals.

[10] Sartono, B., Bodro, D., & Dito, G. (2020). Teknik Eksplorasi Data yang Harus Dikuasai Data Scientist.. Bogor: IPB Press.

[11] Everitt, B. S., S. Landau, M. Leese and D. Stahl (2011). Cluster Analysis. 5th Edition. John Wiley & Sons.

[12] Khofifah, H. N. (2022). Robust Spatial Durbin Model (RSDM) untuk Pemodelan Tingkat Pengangguran Terbuka (TPT) di Provinsi Jawa Barat. *Jurnal Riset Statistika*, *1*(2), 135–142. https://doi.org/10.29313/jrs.v1i2.522

[13] Oktoriandi, D. (2022). Penerapan uji Q Cochran terhadap Atribut Produk Laptop Menggunakan Multiple Response Analysis (MRA). *Jurnal Riset Statistika*, *1*(2), 127–

134. https://doi.org/10.29313/jrs.v1i2.521

[14] Sahwa Chanigia Viqri, Z., & Kurniati, E. (2023). Perbandingan Penerapan Metode Fuzzy Time Series Model Chen-Hsu dan Model Lee dalam Memprediksi Kurs Rupiah terhadap Dolar Amerika. 1(1), 19–26. https://doi.org/10.29313/datamath.v1i1.12