

## Pengelompokan Prioritas Negara Yang Membutuhkan Bantuan Menggunakan Clustering K-Means dengan Elbow dan Silhouette

Yogiek Indra Kurniawan<sup>\*1</sup>, Priandika Ratmadani Anugrah<sup>2</sup>, Rochmat Mulyo Sugihono<sup>3</sup>, Faris Akbar Abimanyu<sup>4</sup>, Lasmedi Afuan<sup>5</sup>

<sup>1,2,3,4,5</sup>Informatika, Fakultas Teknik, Universitas Jenderal Soedirman, Indonesia  
Email: <sup>1</sup>[yogiek@unsoed.ac.id](mailto:yogiek@unsoed.ac.id), <sup>2</sup>[priandika.anugrah@mhs.unsoed.ac.id](mailto:priandika.anugrah@mhs.unsoed.ac.id),  
<sup>3</sup>[rochmat.sugihono@mhs.unsoed.ac.id](mailto:rochmat.sugihono@mhs.unsoed.ac.id), <sup>4</sup>[faris.abimanyu@mhs.unsoed.ac.id](mailto:faris.abimanyu@mhs.unsoed.ac.id),  
<sup>5</sup>[lasmedi.afuan@unsoed.ac.id](mailto:lasmedi.afuan@unsoed.ac.id)

### Abstrak

Sistem kesehatan mencakup personal, lembaga, komoditas, informasi, pembiayaan, dan strategi pemerintah dalam memberikan layanan kesehatan kepada Masyarakat dengan tujuan untuk memenuhi kebutuhan dan harapan masyarakat secara adil dan merata. Status kesehatan masyarakat penting untuk meningkatkan kualitas sumber daya manusia dan produktivitas dari sebuah negara. LSM kemanusiaan HELP International memilih negara-negara yang membutuhkan bantuan berdasarkan faktor sosial, ekonomi, dan kesehatan. Maka dari itu penelitian ini bertujuan melakukan clustering untuk mengelompokkan prioritas negara yang membutuhkan bantuan. Metode yang digunakan dalam pengelompokan negara menggunakan algoritma K-Means dengan metode Elbow dan Silhouette. Tools yang digunakan dalam pengelompokan adalah python. Clustering dan pencarian Silhouette Coefficient dilakukan menggunakan Tools Orange. Dataset yang digunakan mencakup informasi tentang negara-negara di seluruh dunia. Hasil dari penelitian ini adalah clustering negara-negara yang termasuk dalam kelompok C5 hingga C1, dengan kebutuhan prioritas tertinggi di C5 dan terendah di C1.

**Kata kunci:** *clustering, elbow, k-means, negara, python, silhouette*

### *Priority Grouping of Countries That Need Help Using K-Means Clustering with Elbow and Silhouette*

#### *Abstract*

*The health system includes personnel, institutions, commodities, information, financing and government strategies in providing health services to the community with the aim of meeting the needs and expectations of the community fairly and evenly. Public health status is important to improve the quality of human resources and productivity of a country. Humanitarian NGO HELP International selects countries in need of assistance based on social, economic and health factors. Therefore, this research aims to carry out clustering to group the priorities of countries that need assistance. The method used in grouping countries uses the K-Means algorithm with the Elbow and Silhouette methods. The tool used in grouping is Python. Clustering and Silhouette Coefficient search were carried out using Orange Tools. The dataset used includes information about countries around the world. The result of this research is a clustering of countries included in groups C5 to C1, with the highest priority needs in C5 and the lowest in C1.*

**Keywords:** *clustering, country, elbow, k-means, python, silhouette*

## 1. PENDAHULUAN

Sistem kesehatan merupakan suatu istilah yang mencakup personal, lembaga, komoditas, informasi, pembiayaan dan strategi tata pemerintah dalam memberikan layanan pencegahan dan pengobatan kepada masyarakat. Sistem kesehatan dibuat dengan tujuan dapat merespon kebutuhan dan harapan yang dimiliki masyarakat dalam pemenuhan pelayanan kesehatan yang adil dan merata[1][2]. Hal ini tertuang dalam UUD 1945, yang menegaskan bahwa “setiap orang berhak hidup sehat serta berhak memperoleh pelayanan kesehatan”, Pasal 28 H angka (1) “negara bertanggung jawab atas penyediaan fasilitas pelayanan kesehatan bagi seluruh warga negaranya”[3].

Status kesehatan masyarakat di suatu negara sangat penting dalam meningkatkan kualitas sumber daya manusia[4]. Kualitas sumber daya manusia dapat menjadi tolak ukur kemajuan dari negara tersebut. Selain itu,

status kesehatan yang baik juga dapat membuat seseorang menjadi lebih produktif. Hal ini karena kesehatan merupakan hal yang penting untuk menentukan keberlangsungan hidup seseorang. Batasan kesehatan yang diangkat oleh World Health Organization (WHO) yaitu bahwa kesehatan merupakan keadaan sempurna, baik fisik, mental maupun sosial dan tidak hanya bebas dari penyakit dan cacat Sementara itu, menurut UU RI No.36 Tahun 2009 kesehatan merupakan keadaan sehat, baik secara fisik, mental, spiritual maupun sosial yang memungkinkan setiap orang untuk hidup produktif secara sosial maupun ekonomi[5], [6].

Perdagangan internasional merupakan salah satu rangsangan terhadap perkembangan ekonomi. Kegiatan perdagangan internasional khususnya kegiatan ekspor dan impor menjadi terhambat berdasarkan masalah dari masing-masing negara, seperti fluktuasi nilai tukar, konflik politik, ketidakstabilan mata uang, krisis ekonomi, kendala logistik dan transportasi. Padahal kegiatan ekspor dan impor memegang peranan penting dalam kegiatan ekonomi suatu negara[7].

Perdagangan internasional merupakan salah satu rangsangan terhadap perkembangan ekonomi[8]. Kegiatan perdagangan internasional khususnya kegiatan ekspor dan impor menjadi terhambat berdasarkan masalah dari masing-masing negara, seperti fluktuasi nilai tukar, konflik politik, ketidakstabilan mata uang, krisis ekonomi, kendala logistik dan transportasi. Padahal kegiatan ekspor dan impor memegang peranan penting dalam kegiatan ekonomi suatu negara[9].

HELP International adalah LSM (Lembaga Swadaya Masyarakat) kemanusiaan internasional yang berkomitmen untuk memerangi kemiskinan dan menyediakan fasilitas dan bantuan dasar bagi orang-orang di negara-negara terbelakang selama masa bencana dan bencana alam. Jadi, lembaga tersebut harus membuat keputusan untuk memilih negara-negara yang paling membutuhkan bantuan. Karenanya, dibutuhkan pengelompokan negara menggunakan beberapa faktor sosial ekonomi dan kesehatan yang menentukan perkembangan negara secara keseluruhan. Kemudian diputuskan negara-negara yang paling perlu membutuhkan bantuan. Untuk membantu proses prediksi maupun pengelompokan, dapat dipergunakan konsep data mining[10]–[13].

Pengelompokan (*clustering*) adalah sebuah konsep dalam data mining yang berfungsi untuk mengelompokkan data berdasarkan kemiripan karakteristik antar data tersebut[14], [15]. Beberapa Metode yang dapat digunakan untuk *clustering*, antara lain Fuzzy C-Means serta K-Means[16]–[18]. Algoritma K-Means merupakan algoritma *clustering* yang sering digunakan karena kemudahan dalam penggunaan serta ketepatan dalam pengelompokan[19].

Berdasarkan latar belakang diatas, penelitian ini akan melakukan clustering untuk mengelompokkan prioritas negara yang membutuhkan bantuan dengan menggunakan Metode K-Means. Metode Elbow dan Silhouette dipergunakan untuk menentukan jumlah cluster paling ideal dalam pengelompokan pada kasus yang diangkat. Dengan melakukan clustering, diharapkan dapat mengelompokkan negara yang benar-benar lebih membutuhkan bantuan, dan tentunya dapat mengetahui tingkat kebutuhan dari tiap-tiap negara.

## 2. METODE PENELITIAN

Metode yang digunakan untuk pengelompokan negara menggunakan beberapa faktor sosial ekonomi dan kesehatan adalah dengan melakukan Data Mining Clustering Negara. Dalam hal ini, peneliti melakukan Clustering Menggunakan Algoritma K-Means Metode Elbow dan Silhouette.

### 2.1. Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari Kaggle, sebuah platform sumber terbuka untuk dataset. Dataset ini mencakup informasi negara-negara di seluruh dunia dan atribut-atribut yang relevan untuk analisis clustering. Informasi yang terdapat dalam dataset meliputi

- country (Nama negara)
- child (Kematian anak di bawah usia 5 tahun per 1000 kelahiran hidup)
- exports (Ekspor barang dan jasa per kapita. Diberikan sebagai %usia dari PDB per kapita)
- health (Total belanja kesehatan per kapita. Diberikan sebagai %usia dari PDB per kapita)
- imports (Impor barang dan jasa per kapita. Diberikan sebagai %usia dari PDB per kapita)
- income (Pendapatan bersih per orang)
- inflation (Pengukuran tingkat pertumbuhan tahunan Total PDB)
- live\_expec (Rata-rata jumlah tahun seorang anak yang baru lahir akan hidup jika pola kematian saat ini tetap sama)
- total\_fer (Jumlah anak yang akan dilahirkan oleh setiap wanita jika tingkat kesuburan usia saat ini tetap sama)
- gdpp (PDB per kapita. Dihitung sebagai Total PDB dibagi dengan total populasi).

Data ini memberikan gambaran yang komprehensif tentang kondisi sosial, ekonomi, dan kesehatan negara-negara tersebut.

## 2.2. Pemahaman Data

Sebelum dilakukan analisis clustering, dilakukan pemahaman mendalam terhadap dataset yang digunakan. Analisis statistik deskriptif dilakukan untuk mengidentifikasi distribusi data, nilai-nilai yang hilang, serta potensi adanya outlier pada atribut-atribut yang relevan. Dalam konteks ini, atribut country tidak dihitung saat clustering, namun dibutuhkan saat pembuatan kesimpulan bahwa suatu negara termasuk kelompok atau cluster berapa.

## 2.3. Pra-Pemrosesan Data

Pada tahap ini, dilakukan pra-pemrosesan data menggunakan Orange Data Mining untuk mempersiapkan dataset sebelum melakukan analisis clustering. Proses ini meliputi penghapusan nilai-nilai yang hilang, penanganan outlier, dan normalisasi data. Missing values pada dataset diisi menggunakan teknik pengisian seperti mean atau median yang disediakan oleh Orange. Outlier diidentifikasi menggunakan metode deteksi outlier yang telah disediakan oleh software. Selanjutnya, dilakukan normalisasi data untuk mengatasi perbedaan skala pada atribut yang berbeda.

## 2.4. Metode Elbow dan Silhouette untuk Menentukan Jumlah Cluster Optimal

Dalam analisis clustering menggunakan metode K-Means, penting untuk menentukan jumlah cluster yang optimal. Dalam penelitian ini, metode Elbow dan Silhouette digunakan dengan bantuan Python dan Orange Data Mining untuk menentukan jumlah cluster yang paling sesuai dengan dataset.

Metode Elbow merupakan suatu metode yang dapat digunakan untuk menghasilkan informasi dalam menentukan jumlah cluster terbaik dengan cara melihat persentase hasil perbandingan antara jumlah cluster yang akan membentuk siku pada suatu titik tertentu[18]. Metode Elbow Ini memberikan ide atau gagasan dengan cara memilih nilai cluster dan kemudian menambah nilai cluster tersebut untuk dijadikan model data dalam penentuan cluster terbaik[20]. Metode Elbow memperhatikan nilai perbandingan (dari perhitungan SSE untuk setiap nilai cluster) antara jumlah cluster yang akan membentuk siku pada suatu titik, sehingga semakin besar jumlah cluster k maka nilai SSE akan semakin kecil. Rumus SSE sebagai berikut[21].

$$SSE = \sum_{k=1}^k \sum_{x_i} |x_i - C_k|^2 \quad (1)$$

Dimana

- k : cluster ke-c
- $x_i$  : jarak data objek ke-i
- $C_k$  : pusat cluster ke-i

Setelah dilihat akan ada beberapa nilai K yang mengalami penurunan paling besar dan selanjutnya hasil dari nilai K akan turun secara perlahan-lahan sampai hasil dari nilai K tersebut stabil. Hasil persentase yang berbeda dari setiap nilai cluster dapat ditunjukkan dengan menggunakan grafik sebagai sumber informasinya. Jika nilai cluster pertama dengan nilai cluster kedua memberikan sudut dalam grafik atau nilainya mengalami penurunan paling besar maka nilai cluster tersebut yang terbaik[20].

Metode Silhouette merupakan metode evaluasi cluster yang menggabungkan metode cohesion dan separation. Cohesion diukur dengan menghitung seluruh objek yang terdapat dalam sebuah cluster dan separation diukur dengan menghitung jarak rata-rata setiap objek dalam sebuah cluster dengan cluster terdekatnya[22]. Jarak antara data dihitung dengan menggunakan rumus euclidean distance. Untuk menyediakan informasi tentang kualitas hasil clustering pada proses clustering, dapat dihitung silhouette dari masing-masing cluster bahkan keseluruhan cluster dari hasil kerja suatu algoritma clustering. Nilai silhouette untuk keseluruhan data dengan jumlah cluster k, dapat didefinisikan sebagai  $sil(c)$  yang dihitung dengan rumus rata-rata silhouette value untuk semua cluster sebagai berikut[23].

$$sil(c) = sil(k) \frac{1}{|k|} \sum_{i=1}^k sil(c_i) \quad (2)$$

Dimana

- $sil(k)$  : nilai silhouette semua cluster
- $|k|$  : banyaknya cluster k
- $sil(c_i)$  : rata-rata nilai silhouette

Silhouette Coefficient memiliki nilai dengan rentang -1 hingga 1, dengan ketentuan Cluster yang semakin baik ketika nilai dari rata-rata yang dihasilkan Silhouette Coefficient dekat dengan rentang nilai 1 dan sebaliknya, Cluster yang buruk memiliki nilai dari rata-rata yang dihasilkan Silhouette Coefficient dekat dengan rentang nilai -1. Rentang nilai ini dipergunakan untuk menunjukkan korelasi kemiripan dari data yang telah dikelompokkan ke dalam Cluster[24].

Metode elbow menggunakan Python karena ketiadaan metode elbow pada Orange Data Mining. Penggunaan python dalam konteks ini hanya melihat SSE dan differential untuk menentukan jumlah cluster yang sesuai. Sedangkan clustering dan pencarian Silhouette Coefficient menggunakan Orange Data Mining, termasuk melihat plot persebaran tiap negara berdasarkan 2 atribut yang ditentukan.

**2.5. Analisis Clustering dengan K-Means**

Setelah menentukan jumlah cluster optimal menggunakan metode Elbow dan Silhouette, dilakukan analisis clustering dengan menggunakan algoritma K-Means dalam Orange Data Mining. Algoritma K-Means mengelompokkan negara-negara ke dalam kelompok berdasarkan atribut-atribut yang relevan. Proses iteratif dilakukan untuk mengoptimalkan penempatan pusat kelompok. Hasil clustering dievaluasi menggunakan matrik evaluasi seperti Silhouette Coefficient untuk mengukur kualitas dan validitas clustering.

Algoritma K-means adalah algoritma pengelompokan iteratif sederhana. Menggunakan jarak sebagai metrik dan diberi kelas K dalam kumpulan data, hitung rata-rata jarak, berikan pusat massa awal, dengan setiap kelas dijelaskan oleh centroid. Untuk himpunan data X yang berisi n titik data multidimensi dan kategori K yang akan dibagi, jarak Euclidean dipilih sebagai indeks kesamaan dan target pengelompokan meminimalkan jumlah kuadrat dari berbagai jenis; yaitu, meminimalkan.

$$d = \sum_{k=1}^k \sum_{i=1}^n ||(x_i - u_k)||^2 \tag{3}$$

Dimana

- k : pusat gugus K
- $u_k$  : pusat ke-k
- $x_i$  : titik ke-i di himpunan data

Solusi untuk centroid  $u_k$  adalah sebagai berikut:

$$\begin{aligned} \frac{\partial}{\partial u_k} &= \frac{\partial}{\partial u_k} \sum_{k=1}^k \sum_{i=1}^n (x_i - u_k)^2 \\ &= \sum_{k=1}^k \sum_{i=1}^n \frac{\partial}{\partial u_k} (x_i - u_k)^2 \tag{4} \\ &= \sum_{i=1}^n 2(x_i - u_k) \end{aligned}$$

Jadikan persamaan (4) menjadi 0, maka.

$$u_k = \frac{1}{n} \sum_{i=1}^n x_i$$

Ide sentral implementasi algoritma adalah untuk secara acak mengekstraksi titik sampel K dari kumpulan sampel sebagai pusat cluster awal: Bagilah setiap titik sampel ke dalam cluster yang diwakili oleh titik pusat terdekat, maka titik pusat dari semua titik sampel pada setiap klaster adalah titik pusat cluster tersebut. Ulangi langkah diatas hingga titik pusat cluster tidak berubah atau mencapai jumlah iterasi yang ditetapkan. Hasil algoritma berubah dengan pilihan titik pusat, menghasilkan ketidakstabilan hasil. Penentuan titik pusat tergantung pada pemilihan nilai K, yang merupakan fokus dari algoritma; itu secara langsung mempengaruhi hasil clustering, seperti optimalitas lokal atau optimalitas global[25].

**3. HASIL DAN PEMBAHASAN**

Saat akan melakukan analisis clustering, dilakukan pra-pemrosesan dahulu agar tidak ada missing value. Pada penelitian ini, pra-pemrosesan dilakukan pada software Orange, dan missing value pada dataset diisi dengan teknik mean, yaitu rata-rata dari nilai semua row dalam satu atribut. Berikut Tabel 1 yang merupakan data sebelum dilakukan pra-pemrosesan, dan Tabel 2 yang merupakan data setelah dilakukan pra-pemrosesan.

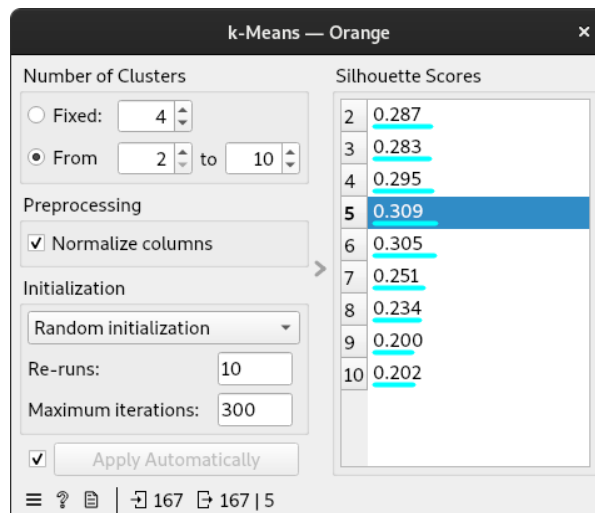
Tabel 1. Data Sebelum Pra-pemrosesan

country	child_mort	exports	health	...	gdp
Afganistan	90.2	10.000	7.58	...	553
Albania	16.6	28.000	6.55	...	4090
Algeria	27.3	38.400	4.17	...	4460
Angola	119.0	62.300	2.85	...	3530
...	...	...	...	...	...
Zambia	83.1	37.000	5.89	...	1460

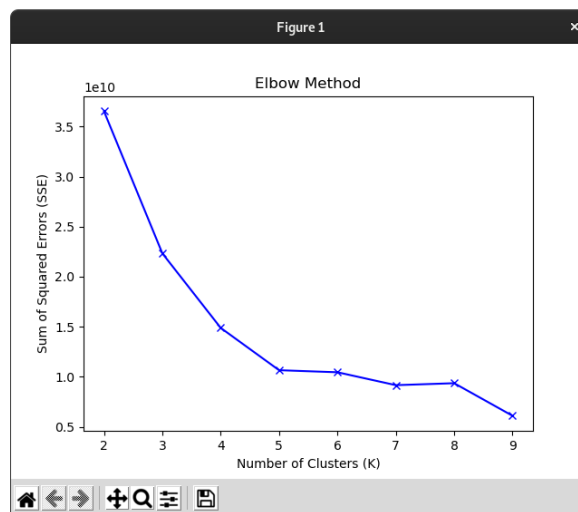
Tabel 2. Data Setelah Pra-pemrosesan

country	child_mort	exports	health	...	gdp
Afganistan	90.2	10.000	7.58	...	553
Albania	16.6	28.000	6.55	...	4090
Algeria	27.3	38.400	4.17	...	4460
Angola	119.0	62.300	2.85	...	3530
...	...	...	...	...	...
Zambia	83.1	37.000	5.89	...	1460

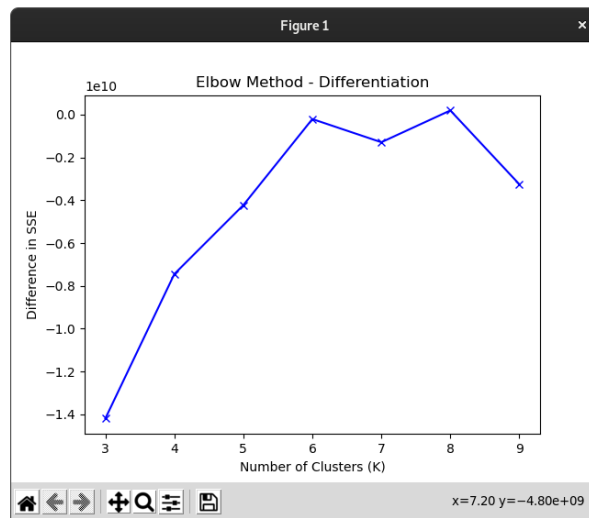
Untuk menentukan jumlah cluster terbaik menggunakan metode elbow dan silhouette. Pencarian nilai SSE menggunakan Python, sedangkan nilai silhouette menggunakan software Orange.



Gambar 3. Nilai Silhouette Tiap Jumlah Cluster



Gambar 4. Nilai SSE (Sum of Squared Errors) Setiap Jumlah Cluster



Gambar 5. Nilai Differentiation SSE Setiap Jumlah Cluster

Dari hasil metode silhouette, didapatkan jumlah cluster terbaik yaitu 5, dengan nilai silhouette 0.309. Kemudian pada metode elbow, grafik nilai SSE tiap cluster yang lebih membentuk siku pada suatu titik yaitu ketika nilai K adalah 5. Jadi jumlah cluster yang optimal adalah sebanyak 5 cluster.

Tabel 3. Perbandingan menggunakan metode elbow dan silhouette

Jumlah Cluster	Metode	
	Elbow	Silhouette
2	36528387934.3221	0.287
3	22336241763.9737	0.283
4	14891567194.7864	0.295
5	10956489943.59941	0.309
6	10486940156.596891	0.305
7	9495804440.307653	0.251
8	9057462129.02651	0.234
9	6330834153.916922	0.200
10	5815483553.063592	0.202

Masing-masing metode memiliki kelebihan dan kekurangannya, maka perlu ketepatan dalam memadukan metode clustering yang digunakan, metode untuk menentukan jumlah cluster yang tepat dan struktur data serta ukuran data dengan memaksimalkan kelebihan setiap metode yang ada[26]. Metode Elbow dan Silhouette digunakan untuk menemukan jumlah cluster yang optimal. Ketidakjelasan muncul untuk metode siku untuk mengambil nilai k. Analisis siluet dapat digunakan untuk mempelajari jarak pemisahan antara cluster yang dihasilkan dan dapat dianggap sebagai metode yang lebih baik dibandingkan dengan metode Elbow[27].

Setelah menjalankan analisis clustering menggunakan metode K-Means dengan jumlah cluster optimal yang ditentukan pada aplikasi Orange, didapatkan hasil pengelompokan negara-negara ke dalam kelompok berdasarkan tingkat kebutuhan bantuan. Berikut adalah Tabel 3 yang merupakan centroid tiap atribut tiap cluster, dan Tabel 4 yang merupakan hasil dari clustering.

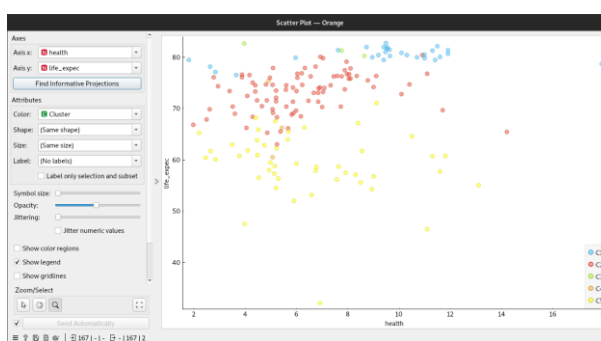
Tabel 3. Centroid Tiap Atribut Tiap Cluster

Cluster	Silhouette	child_mort	exports	...	gdpp
C1	0.591439	-0.82788	0.168112	...	1.701970
C2	0.602131	-0.42697	0.009869	...	-0.318896
C3	0.618071	-0.84900	4.935673	...	2.440797
C4	0.494732	0.48406	-0.278413	...	-0.372346
C5	0.585144	1.30630	-0.418849	...	-0.605453

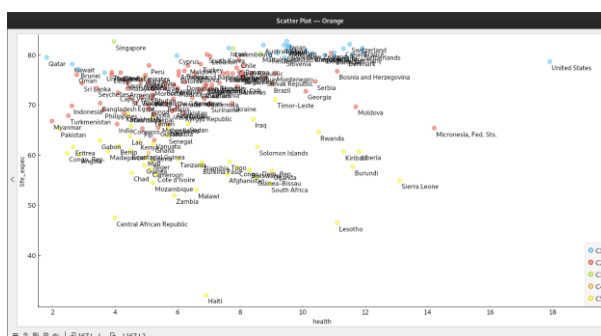
Tabel 4. Data Hasil Clustering

country	Cluster	Silhouette
Afganistan	C5	0.638269
Albania	C2	0.650317
Algeria	C2	0.620313
Angola	C5	0.606156
...	...	...
Zambia	C5	0.632653

Dan berikut adalah plot persebaran negara berdasarkan nilai variabel yang ditentukan, seperti contoh yaitu health dan life expect.



Gambar 6. Plot Persebaran Negara Berdasarkan Health dan Life Expec Hanya Menggunakan Penanda Cluster



Gambar 7. Plot Persebaran Negara Berdasarkan Health dan Life Expec Menggunakan Penanda Nama Negara

Berikut adalah Interpretasi Hasil Clustering:

1. C1 (Tingkat Kebutuhan Bantuan Sangat Rendah)  
Negara ini memiliki tingkat pembangunan manusia yang sangat tinggi, tingkat kemiskinan yang rendah, dan angka harapan hidup yang tinggi. Mereka merupakan negara dengan kondisi sosial ekonomi yang sangat stabil dan memerlukan bantuan yang sangat terbatas.
2. C2 (Tingkat Kebutuhan Bantuan Rendah)  
Negara ini memiliki tingkat pembangunan manusia yang tinggi, tingkat kemiskinan yang rendah, dan angka harapan hidup yang tinggi. Mereka memiliki kondisi sosial ekonomi yang stabil dan memerlukan bantuan yang lebih sedikit dibandingkan dengan kelompok sebelumnya.
3. C3 (Tingkat Kebutuhan Bantuan Sedang)  
Negara ini memiliki tingkat kemiskinan yang lebih rendah dibandingkan dengan Kelompok 1 dan 2, serta HDI yang sedang. Namun, mereka masih membutuhkan bantuan dalam beberapa aspek seperti pengembangan infrastruktur dan akses ke layanan dasar. Negara ini memiliki tingkat kemiskinan yang lebih rendah lagi, HDI yang sedang, dan tingkat harapan hidup yang lebih tinggi. Mereka mungkin membutuhkan bantuan terfokus dalam bidang-bidang seperti pendidikan, pelatihan, dan pemberdayaan ekonomi.
4. C4 (Tingkat Kebutuhan Bantuan Tinggi)  
Negara ini menghadapi tingkat kemiskinan yang tinggi, HDI yang rendah, dan angka harapan hidup yang relatif rendah. Meskipun tidak seburuk Kelompok 1, negara ini memerlukan bantuan yang cukup signifikan dalam mengatasi tantangan sosial ekonomi.

5. C5 (Tingkat Kebutuhan Bantuan Sangat Tinggi)  
Negara ini memiliki tingkat kemiskinan yang sangat tinggi, indeks pembangunan manusia (HDI) yang rendah, angka harapan hidup yang rendah, serta masalah kesehatan dan pendidikan yang serius. Negara ini membutuhkan bantuan signifikan dalam berbagai aspek.  
Dari hasil clustering tersebut, negara yang menjadi prioritas untuk penyaluran bantuan, yaitu dimulai dari C5.

#### 4. KESIMPULAN

Berdasarkan hasil penelitian tersebut dapat diambil kesimpulan, prioritas negara yang membutuhkan bantuan dimulai dari negara yang termasuk kelompok C5, sampai dengan negara yang termasuk kelompok C1. Berdasarkan hasil clustering negara-negara dalam kelompok yang termasuk tingkat kebutuhannya sangat rendah(C1) di antaranya Brunei, Qatar, dan Uni Arab Emirates, negara-negara dalam kelompok yang termasuk tingkat kebutuhannya rendah(C2) di antaranya Rusia, China, dan Indonesia, negara-negara dalam kelompok yang termasuk tingkat kebutuhannya sedang(C3) di antaranya Singapore, Malta, dan Luxembourg, negara-negara dalam kelompok yang termasuk tingkat kebutuhannya tinggi(C4) di antaranya Mongolia, Nigeria, dan Venezuela, dan negara-negara dalam kelompok yang termasuk tingkat kebutuhannya sangat tinggi (C5) di antaranya Guinea, Central African Republic, dan Kenya.

#### DAFTAR PUSTAKA

- [1] A. Hanifa, Y. I. Kurniawan, J. H. Husein, A. K. Nugroho, and I. Permadi, "Prediction of patient length of stay using random forest method based on the Indonesian national health insurance," *J. Infotel Informatics - Telecommun. - Electron.*, vol. 15, no. 3, pp. 233–240, 2023, doi: <https://doi.org/10.20895/infotel.v15i3.963>.
- [2] A. A. Rahman and Y. I. Kurniawan, "Aplikasi Klasifikasi Penerima Kartu Indonesia Sehat Menggunakan," *Progr. Stud. Inform. Univ. Muhammadiyah Surakarta*, 2016.
- [3] R. N. Putri, "Perbandingan Sistem Kesehatan di Negara Berkembang dan Negara Maju," *J. Ilm. Univ. Batanghari Jambi*, vol. 19, no. 1, pp. 139–145, 2019, doi: 10.33087/jiubj.v19i1.572.
- [4] Y. I. Kurniawan and W. Dwiyatmika, "Aplikasi diagnosa retardasi mental pada anak," in *Prosiding SEMNAS Penguatan Individu di Era Revolusi Informasi*, 2017, pp. 336–343, [Online]. Available: <https://publikasiilmiah.ums.ac.id/handle/11617/9053>.
- [5] L. S. Rakasiwi and A. Kautsar, "Pengaruh Faktor Demografi dan Sosial Ekonomi terhadap Status Kesehatan Individu di Indonesia," *Kaji. Ekon. Dan Keuang.*, vol. 5, no. 2, pp. 146–157, 2021.
- [6] K. K. R. Indonesia, "Undang-Undang RI tentang Kesehatan 2009," 2009. [https://infeksiemerging.kemkes.go.id/download/UU\\_36\\_2009\\_Kesehatan.pdf](https://infeksiemerging.kemkes.go.id/download/UU_36_2009_Kesehatan.pdf).
- [7] U. Zakia, S. Anwar, and I. Y. Ulya, "DAMPAK PANDEMI COVID-19 TERHADAP IMPOR PROVINSI ACEH BERDASARKAN HASIL CLUSTERING NEGARA ASAL IMPOR," *Bul. Ilm. Litbang Perdagangan.*, vol. 16, no. 2, 2022, doi: 10.55981/bilp.2022.3.
- [8] N. B. Rahmawan and S. I. Oktora, "The impact of zero import tariff policy and air pollution prevention and control action plan on Indonesian coal export to China," *Bul. Ilm. Litbang Perdagangan.*, vol. 12, no. 1, pp. 73–94, 2018.
- [9] I. R. Warer, N. Putu, and W. Setyari, "PENGARUH EKSPOR MIGAS, PENANAMAN MODAL ASING, UTANG LUAR NEGERI, DAN INFLASI TERHADAP PERTUMBUHAN EKONOMI INDONESIA," *J. Ekon. DAN BISNIS*, vol. 10, no. 12, pp. 1063–1076, 2021.
- [10] Y. I. Kurniawan, A. Fatikasari, M. L. Hidayat, and M. Waluyo, "Prediction For Cooperative Credit Eligibility Using Data Mining Classification With C4.5 Algorithm," *J. Tek. Inform.*, vol. 2, no. 2, pp. 57–64, 2021, doi: 10.20884/1.jutif.2021.2.2.49.
- [11] N. R. Indraswari and Y. I. Kurniawan, "Aplikasi Prediksi Usia Kelahiran Dengan Metode Naive Bayes," *J. Simetris*, vol. 9, no. 1, pp. 129–138, 2018, doi: 10.24176/simet.v9i1.1827.
- [12] D. Febriyanto and Y. I. Kurniawan, "Prediksi Penyakit Tuberculosis (TBC) Menggunakan Algoritma C4.5," *J. Ilm. SINUS*, vol. 16, no. 2, pp. 23–36, 2018, doi: 10.30646/sinus.v16i2.366.
- [13] D. A. Kurniawan and Y. I. Kurniawan, "Aplikasi Prediksi Kelayakan Calon Anggota Kredit Menggunakan Algoritma Naïve Bayes," *J. Teknol. dan Manaj. Inform.*, vol. 4, no. 1, pp. 1–12, 2018, doi: 10.26905/jtmi.v4i1.1831.



- 
- [14] K. Peng, V. C. M. Leung, and Q. Huang, "Clustering Approach Based on Mini Batch Kmeans for Intrusion Detection System over Big Data," *IEEE Access*, vol. 6, pp. 11897–11906, 2018, doi: 10.1109/ACCESS.2018.2810267.
- [15] Y. S. Nugroho, S. Islam, D. Gunawan, Y. I. Kurniawan, and M. J. Hossain, "Dataset of network simulator related-question posts in stack overflow," *Data Br.*, vol. 41, no. 107942, pp. 1–7, 2022, doi: 10.1016/j.dib.2022.107942.
- [16] A. V. Efrilla, S. B. Sulisty, K. Wijaya, P. H. Kuncoro, and A. Sudarmaji, "Klasifikasi Penyakit Pada Daun Stroberi Menggunakan K-Means Clustering dan Jaringan Syaraf Tiruan," *J. Keteknikan Pertan. Trop. dan Biosist.*, vol. 8, no. 2, pp. 161–170, 2020, doi: 10.21776/ub.jkptb.2020.008.02.06.
- [17] Taslim, D. Toresa, D. Jollyta, D. Suryani, and E. Sabna, "Optimasi K-Means dengan Algoritma Genetika untuk Target Pemanfaat Air Bersih Provinsi Riau," *Indones. J. Comput. Sci.*, vol. 10, no. 1, pp. 1–12, 2021.
- [18] N. H. Harani, C. Prianto, and F. A. Nugraha, "egmentasi Pelanggan Produk Digital Service Indihome Menggunakan Algoritma K-Means Berbasis Python," *J. Manaj. Inform.*, vol. 10, no. 2, pp. 133–146, 2020.
- [19] M. A. Ghofar and Y. I. Kurniawan, "APLIKASI PENGELOMPOKAN PELANGGAN PADA UMS STORE MENGGUNAKAN ALGORITMA K-MEANS," *J. Teknol. Manaj. Inform.*, vol. 4, no. 1, pp. 196–203, 2018.
- [20] E. Muningsih and S. Kiswati, "Sistem aplikasi berbasis optimasi metode elbow untuk penentuan clustering pelanggan," *Joutica J. Inform. Unisla*, vol. 3, no. 1, pp. 117–124, 2018.
- [21] N. T. Hartanti, "Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional," *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 2, pp. 82–89, 2020, doi: 10.25077/teknosi.v6i2.2020.82-89.
- [22] B. Wira, A. E. Budianto, and A. S. Wiguna, "Implementasi Metode K-Medoids Clustering Untuk Mengetahui Pola Pemilihan Program Studi Mahasiswa Baru Tahun 2018 Di Universitas Kanjuruhan Malang," *Rainstek J. Terap. Sains Teknol.*, vol. 1, no. 3, pp. 53–68, 2019.
- [23] S. Paembonan and H. Abduh, "Penerapan Metode Silhouette Coefficient untuk Evaluasi Clustering Obat," *PENA Tek. J. Ilm. Ilmu-Ilmu Tek.*, vol. 6, no. 2, pp. 48–54, 2021.
- [24] D. S. Anggraeni and U. Enri, "Pengelompokan Data Kemiskinan Provinsi Jawa Barat Menggunakan Algoritma K-Means dengan Silhouette Coefficient," *TEMATIK*, vol. 9, no. 1, pp. 29–35, 2022.
- [25] C. Yuan and H. Yang, "Research on K-Value Selection Method of K-Means Clustering Algorithm," *J Basel*, vol. 2, no. 2, pp. 226–235, 2019, doi: 10.3390/j2020016.
- [26] D. A. I. C. Dewi and D. A. K. Pramita, "Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali," *Matrix J. Manaj. Teknol. dan Inform.*, vol. 9, no. 3, pp. 102–109, 2019.
- [27] Ichi.Pro, "Metode Silhouette - Lebih Baik daripada Metode Siku untuk menemukan Cluster Optimal," 2023. <https://ichi.pro/id/metode-silhouette-lebih-baik-daripada-metode-siku-untuk-menemukan-cluster-optimal-61080390822033> (accessed Jun. 17, 2023).