

## ANALISIS CLUSTER NEGARA DI ASIA BERDASARKAN TINGKAT KENYAMANAN HIDUP MENGGUNAKAN METODE K-MEANS

Muhammad Muhtadin Billah, Daud Rasyid Al-Hadi, Dina Zatusiva Haq, Dian C.R. Novitasari

Matematika, Universitas Islam Negeri Sunan Ampel Surabaya

Jl. Dr. Ir. H. Soekarno No.682, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur 60294

diancrininov@gmail.com

### ABSTRAK

Kenyamanan hidup adalah konsep penting yang sering dijadikan tolok ukur kualitas hidup individu dan masyarakat. Berbagai faktor seperti Crime Index, Safety Index, GDP, Air Quality, dan Education Quality digunakan untuk menentukan tingkat kenyamanan hidup suatu negara. Negara-negara di Asia bersaing untuk meningkatkan tingkat kenyamanan hidup guna menarik warga asing dan memperbaiki kualitas negara, namun belum banyak penelitian yang menggunakan metode K-Means untuk menganalisis tingkat kenyamanan hidup di negara Asia. Penelitian ini bertujuan untuk mengelompokkan data negara-negara di Asia berdasarkan tingkat kenyamanan hidup menggunakan metode K-Means clustering. Algoritma K-Means digunakan untuk membagi data menjadi beberapa cluster berdasarkan kesamaan karakteristik, melalui analisis data, penentuan jumlah cluster, penentuan centroid, analisis jarak Euclidean, pengelompokkan data berdasarkan jarak minimum, dan evaluasi cluster menggunakan silhouette index. Hasil penelitian menunjukkan bahwa dari 45 data yang dianalisis, negara-negara di Asia dapat dikelompokkan menjadi tiga cluster tingkat kenyamanan hidup: rendah, sedang, dan tinggi, dengan 32 negara masuk dalam cluster rendah, 2 negara dalam cluster sedang, dan 11 negara dalam cluster tinggi. Evaluasi menggunakan silhouette index menghasilkan nilai sebesar 0,761, menunjukkan tingkat akurasi yang tinggi. Penelitian ini diharapkan dapat memberikan gambaran mengenai tingkat kenyamanan hidup di negara-negara Asia dan menjadi acuan bagi pemimpin negara dalam meningkatkan kualitas hidup masyarakat.

**Kata kunci :** tingkat kenyamanan hidup; K-Means; pengelompokkan; negara asia; analisis data

### 1. PENDAHULUAN

Kenyamanan hidup adalah poin yang paling penting dalam pembangunan manusia dan telah dijadikan sebagai tolak ukur kesejahteraan masyarakat. Dalam mencapai kehidupan yang baik tersebut, berbagai tindakan akan dilakukan demi meningkatkan tingkat kesejahteraan/kenyamanan hidup [1].

Pada zaman globalisasi sekarang, negara di Asia terus bersaing untuk mendapatkan nilai kenyamanan yang paling tinggi, agar dapat menarik warga asing sebagai pendapatan dan memperbaiki kualitas negara. Sehingga banyak peneliti yang ingin melakukan riset untuk menganalisis dan membandingkan tingkat kenyamanan hidup di tiap negara di Asia. Menggunakan faktor seperti Crime Index, Safety Index, GDP, Air Quality, dan Education Quality bisa diketahui tingkat dari kenyamanan negara tersebut. Berbagai metode telah dilakukan oleh beberapa ilmuwan untuk menganalisis tingkat kenyamanan hidup, seperti regresi, dan cluster [2].

Tetapi, metode K-Means belum banyak digunakan untuk menganalisis tingkat kenyamanan hidup di negara Asia.

Metode dalam penelitian ini menggunakan K-Means untuk mengetahui tingkat kenyamanan hidup di tiap negara di Asia. K-Means adalah metode pada sebuah algoritma yang mengelompokkan sebuah data menjadi beberapa cluster dengan memperhatikan keakuratan tiap cluster terhadap data [3].

Tujuan utama analisis cluster adalah untuk membandingkan suatu objek berdasarkan kesamaan karakteristik [4].

Ketika menggunakan metode ini kita harus memperhatikan variabel-variabel yang penting karena penambahan atau pengurangan variabel akan memengaruhi terhadap hasil cluster.

Metode ini memiliki tingkat keakuratan yang tinggi, sehingga dirasa efisien untuk digunakan dalam mengolah data yang berjumlah besar seperti dalam penelitian ini. Penelitian ini membahas tentang cara mengelompokkan data dengan tujuan menganalisis, mengidentifikasi, dan membandingkan Tingkat Kenyamanan Hidup di tiap negara Asia. Dan tujuan adanya jurnal ini adalah untuk mengelompokkan data di tiap negara Asia menggunakan metode K-Means berdasarkan beberapa faktor. Penelitian ini juga memiliki batasan masalah, beberapa batasan masalah yang terdapat didalam penelitian ini adalah, hanya membahas Tingkat Kenyamanan Hidup, daerah yang di cluster hanya negara di Asia, hanya beberapa faktor yang digunakan dalam penelitian Analisis Tingkat Kenyamanan Hidup, hanya metode K-Means yang digunakan dalam sistem pemetaan negara di Asia.

Bisa dilihat dari latar belakang di atas, maka pada penelitian ini digunakan Algoritma Clustering K-Means untuk mengetahui tingkat kenyamanan hidup yang diinginkan beberapa orang, yang terutama dikawasan Asia berdasarkan beberapa faktor. Dengan dibedakan menjadi tiga cluster yaitu tinggi, sedang, dan rendah untuk mengetahui hasil dari penelitian ini. Hasil dari penelitian ini bisa dijadikan sebagai

gambaran tentang tingkat kenyamanan hidup. Diharapkan penelitian ini bisa dijadikan sebagai acuan bagi para pemimpin di tiap negara Asia untuk mempertimbangkan pembangunan negara yang tepat untuk meningkatkan kenyamanan dan kualitas hidup masyarakat. Dan semoga dengan penelitian ini bermanfaat bagi para peneliti dalam mengembangkan penelitian lebih lanjut kedepannya, terutama dibidang kenyamanan hidup masyarakat di negara Asia, dan juga diharapkan adanya penelitian ini masyarakat umum memahami tingkat kenyamanan hidup di negara Asia.

## 2. TINJAUAN PUSTAKA

Berdasarkan penelitian sebelumnya metode K-Means dapat digunakan di berbagai data, seperti dalam pengelompokan kasus Covid-19 di Indonesia, hasilnya menunjukkan nilai SI dengan tingkat akurasi sebesar 0,857 atau 85,7%, artinya tingkat keakurasian data mendekati nilai sebenarnya [5].

Selanjutnya dalam clustering tingkat kemiskinan, hasil penelitian terdapat 3 cluster provinsi di Indonesia berdasarkan tingkat kemiskinannya yaitu Cluster 0 (provinsi dengan tingkat kemiskinan rendah), Cluster 1 (provinsi tingkat kemiskinan sedang), dan Cluster 2 (provinsi dengan tingkat kemiskinan tinggi). Provinsi yang termasuk dalam kategori provinsi dengan tingkat kemiskinan tinggi yaitu Maluku, Papua Barat dan Papua [6].

### 2.1. Kenyamanan Tinggal

Dalam kehidupan, terkhususnya manusia sebagai makhluk hidup. Faktor kenyamanan ialah hal yang sangat penting sehingga perlu diperhatikan secara serius. Salah satu faktor yang memengaruhi perilaku manusia dan menjadi pengendali lingkungan terkait kondisi fisiologis adalah dampak dari adanya kenyamanan [7].

Oleh karena itu, kenyamanan ialah suatu hal yang menjadi dasar atau komponen utama di lingkungan tempat manusia melangsungkan kehidupan. Kondisi setiap manusia di lingkungannya selalu memengaruhi bagaimana tingkat kenyamanan yang akan dirasakan [8].

Ada beberapa faktor yang mempengaruhi seperti, Crime Index, Safety Index, GDP, Air Quality, dan Education Quality.

### 2.2. K-Means Clustering

Metode K-Means Clustering adalah sebuah metode dalam mengolah data mining yang digunakan untuk membagi data-data menjadi berkelompok atau pada klusternya masing-masing [9].

Sedangkan cluster adalah kelompok yang di dalamnya berisi data-data similar atau bisa disebut data tersebut memiliki kesamaan data yang tinggi sehingga bisa dikelompokkan [10].

Dalam satu kelompok/cluster, data di dalamnya memiliki kemiripan yang hampir sama. Kemiripan

tersebut bukan hanya antar data pada satu cluster, melainkan data antar cluster lain pula [11].

Kedekatan yang ada pada satu cluster diambil dari perhitungan jarak ke data dan jarak ke centroid. Adapun algoritma K-Means yaitu mengelompokkan data yang terdapat kemiripan antar data atau bisa disebut punya karakteristik yang sama [12].

Menghitung nilai K-Means yang pertama, menentukan jumlah kelompok berdasarkan data yang akan dibagi. Selanjutnya distribusikan secara acak data ke dalam kelompok atau cluster untuk menentukan pusat centroid. Lalu menghitung centroid dari tiap kelompok. Letak centroid tiap kelompok diambil dari rata-rata semua data tiap variabel. Setelah itu, pindahkan masing-masing data ke centroid terdekat. Caranya menggunakan rumus Euclidean. Selanjutnya mengelompokkan data untuk mengetahui anggota kelompok, dengan cara perhitungan jarak terdekat (minimum) dari sebuah objek data. Kemudian mengulang langkah sebelumnya, jika terdapat data yang kelompok (cluster) yang berbeda sama sebelumnya, maka lakukan langkah yang sama secara berulang hingga mendapatkan nilai centroid tetap dan anggota dari tiap kelompok tidak berubah. Berikutnya jika sudah tidak ada anggota kelompok yang berubah, selanjutnya melakukan evaluasi cluster dengan menggunakan Silhouette Index yang bertujuan untuk mengetahui kesamaan jarak dari tiap cluster. Kemudian mengclusterkan ulang data berdasarkan hasil Silhouette Index tertinggi [13]. Yang terakhir menyimpulkan hasil dari penelitian.

Adapun rumus yang digunakan seperti di bawah ini: [14]

$$d_{(x,y)} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad [1]$$

Keterangan:

d = jarak antara x dan y

x = data pusat cluster

y = data pada atribut

i = setiap data

n = jumlah data

$x_i$  = data pada pusat cluster ke-i

$y_i$  = data pada setiap data ke-i

Saat menghitung jarak matriks, dapat menduga berapa klaster yang ingin kita buat dengan contoh adalah 3 klaster maka  $k=3$ . Tidak lupa pula penentuan centroid yang dihitung dengan rumus sebagai berikut:

$$C_i = \frac{\sum_{j=1}^n x_j \in s_j}{n} \quad [2]$$

Keterangan:

$C_i$  = centroid baru ke-i

$s_i$  = objek ke-i

$x_i$  = nilai pada objek ke-i

n = jumlah data pada tiap kelompok

Jika jarak matriks sudah diketahui, centroid juga sudah di tentukan. Maka dapat melakukan evaluasi silhouette untuk mengevaluasi kluster.

**2.3. Silhouette Index**

Silhouette index adalah sebuah cara yang digunakan untuk memvalidasi sebuah hasil kluster pada dari sebuah data. Metode ini dapat memberikan hasil tentang seberapa baik data setelah dilakukannya uji K-Means [15].

Adapun cara menganalisisnya adalah dengan melihat nilai silhouette nya apabila >1 maka ia tidak berada di kluster yang tidak tepat namun sebaliknya jika nilai silhouette nya <0 maka ia ada di kluster yang tepat [16].

Untuk mendapatkan nilai dari silhouette itu, dapat dilakukan dengan menghitung menggunakan rumus berikut: [17]

$$S_i = \frac{1}{n} \sum_{i=1}^n \left( \frac{b(i)-a(i)}{\max(a(i),b(i))} \right) \quad [3]$$

Keterangan:

$S_i$  = Silhouette Indeks

$a_i$  = jarak rata-rata sampel  $i$  ke sampel lain dalam kluster

$b_i$  = jarak minimum sampel dari sampel  $i$  ke kluster lain

Adapun tabel interpretasi Silhouette Score sebagai berikut:

Tabel 1. Indeks Silhouette Score

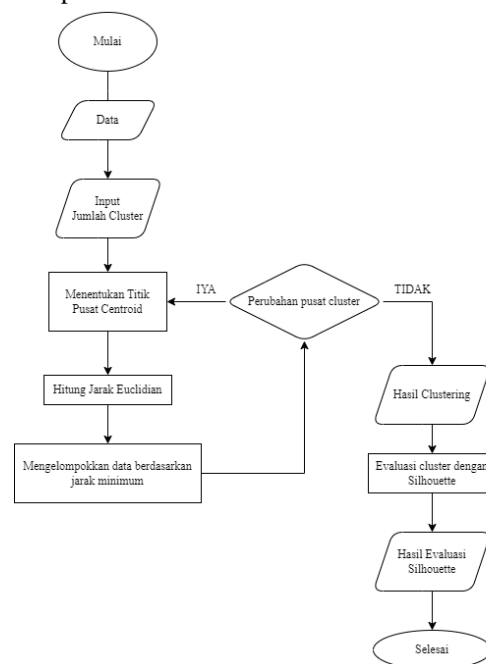
NO	Indeks	Keterangan
1	0,71 - 1	Sangat Baik
2	0,51 - 0,7	Baik
3	0,26 - 0,5	Buruk
4	< 0,25	Sangat Buruk

**3. METODE PENELITIAN**

Data yang digunakan dalam analisis ini adalah data dari negara negara di Asia. Dalam dataset tersebut, negara-negara memiliki bebrapa faktor yang mempengaruhi penelitian ini yakni menganalisis dan melakukan klustering terhadap negara di Asia berdasarkan faktor tingkat kelayakan tinggal suatu negara. Pada data, berisi variabel numeric yang dibagi menjadi beberapa variabel diantaranya crime index, safety index, GDP, air quality, dan education quality. Metode Clustering memiliki beberapa jenis dan salah satunya yang digunakan dalam penelitian ini yaitu K-Means. Adapun prinsip dari metode ini adalah menyusun K buah partisi/pusat massa (centroid)/rata-rata (mean) dari kumpulan data yang di olah datanya menggunakan metode K-Means. Selain itu, tujuan dari

metode ini adalah melakukan pengelompokan data adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan dan memaksimalkan variasi antar kelompok. Dalam penelitian ini menggunakan data yang mempengaruhi tingkat kenyamanan hidup dari tiap negara di Asia, yang akan dikelompokkan menjadi tiga (3). Variabel tersebut akan diteliti untuk mengetahui perbedaan dari tiap negara. Hal ini merupakan langkah awal dalam penelitian ini.

Seluruh data yang digunakan dalam penelitian ini diambil dari internet. Cara pengelompokan data dengan metode K-Means. Dalam melakukan olah data dengan metode K-Means, pertama yang harus dilakukan adalah input data, menentukan titik pusat centroid dihitung menggunakan rumus di persamaan (2), menghitung jarak Euclidean dihitung menggunakan rumus persamaan (1), mengelompokan data berdasarkan jarak minimum, perubahan pusat cluster. Pada perubahan pusat cluster, apabila data sudah tidak berubah maka akan lanjut pada output hasil clustering. Namun apabila masih berubah, maka akan terus melakukan looping sampai data tidak berubah. Setelah mendapatkan hasil clustering, lanjut pada evaluasi dengan silhouette dihitung menggunakan rumus di persamaan (3), dan akan menghasilkan hasil evaluasi yang berguna untuk mengetahui apakah klastering tersebut valid dengan melihat nilai silhouette score. Berikut adalah flowchart dari penelitian ini.



Gambar 1. Flowchart

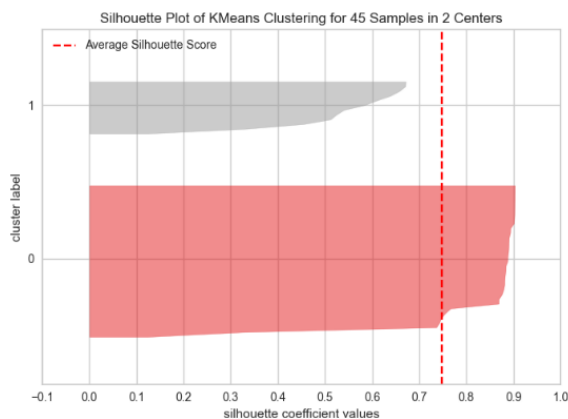
Tabel 2. Data negara di Asia

NO	Country	Crime Index (V1)	Safety Index (V2)	GDP (V3)	Air Quality (V4)	Education Quality (V5)
1.	Afghanistan	78,3	21,7	408	58,8	0,21
2.	Armenia	22,4	77,6	5116	26,4	0,72
3.	Azerbaijan	31,6	68,4	5590	18,8	0,7
4.	Bahrain	23,9	76,1	24395	39,2	0,718
5.	Bangladesh	61,5	38,5	1785	79,9	0,371
...	...	...	...	...	...	...
43.	Uzbekistan	30,5	69,5	3473	28,6	0,739
44.	Vietnam	42,2	57,8	3655	29,6	0,541
45.	Yemen	68,6	31,4	1018	50,4	0,194

4. HASIL DAN PEMBAHASAN

Tujuan dari dilakukannya penelitian ini adalah untuk mengelompokkan atau melakukan clustering terhadap data guna mengetahui tingkat kenyamanan hidup di tiap negara wilayah Asia menggunakan metode K-Means. Data sekunder yang respondennya berasal dari seluruh penduduk di Asia digunakan pada penelitian ini. Adapun beberapa variabel yang menjadi faktor untuk melakukan penelitian ini berjumlah 5 yaitu tingkat kejahatan, tingkat keamanan, GDP, kualitas udara, dan tingkat pendidikan. Jumlah data yang akan digunakan sebanyak 46 data negara di wilayah Asia.

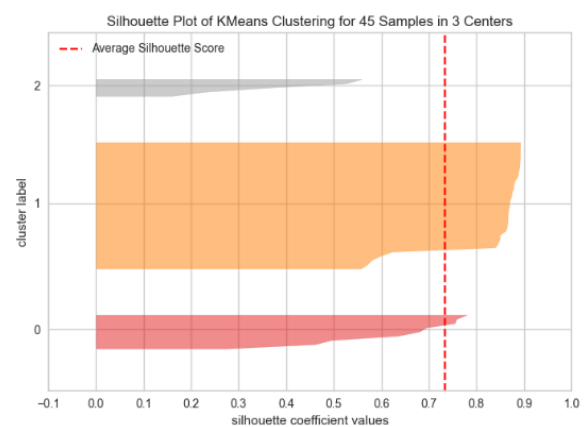
Pada penelitian ini, menggunakan dua model visualisasi yaitu Silhouette Plot. Hasil silhouette umumnya memiliki nilai yang berkisar antara -1 sampai 1. jika nilai hasil silhouette mendekati 1, maka dapat dikatakan bahwa semakin akurat suatu cluster dalam pengelompokan data. Sebaliknya jika nilai hasil silhouette mendekati -1, maka dapat dikatakan bahwa semakin tidak akurat suatu cluster dalam pengelompokan data.



Gambar 2. Silhouette 2 cluster

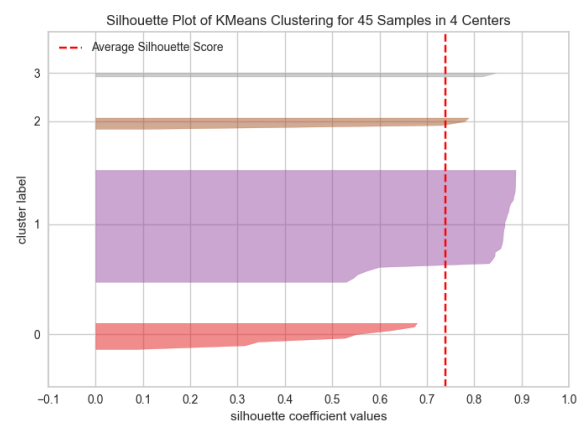
Berdasarkan gambar 2 hasil evaluasi cluster dengan silhouette diatas dapat diketahui bahwa terdapat 2 cluster, pada warna merah adalah cluster 0 dan warna abu adalah cluster 1. Pada hasil silhouette memiliki nilai rata-rata sebesar 0,748. Dari seluruh total data, dilakukan evaluasi silhouette dengan ketentuan 2 cluster seperti gambar di atas. Pada gambar, menunjukkan bahwa banyak dari data masuk

pada cluster 0 dengan nilai rata-rata silhouette pada interval 0,7 hingga 0,8.



Gambar 3. Silhouette 3 cluster

Berdasarkan gambar 3 hasil evaluasi cluster dengan silhouette diatas dapat diketahui bahwa terdapat 3 cluster, pada warna merah adalah cluster 0, warna orange adalah cluster 1, dan warna abu adalah cluster 2. Pada hasil silhouette memiliki nilai rata-rata sebesar 0,761. Dari seluruh total data, dilakukan evaluasi silhouette dengan ketentuan 3 cluster. Setelah dilakukan evaluasi silhouette, hasil menunjukkan bahwa terdapat banyak data yang masuk pada cluster 1 sehingga menjadi kelompok dengan jumlah terbanyak. Selain itu, kelompok dengan jumlah paling sedikit adalah cluster 2.



Gambar 4. Silhouette 4 cluster

Berdasarkan gambar 4, hasil evaluasi cluster dengan silhouette diatas dapat diketahui bahwa terdapat 3 cluster, pada warna merah adalah cluster 0, warna ungu adalah cluster 1, warna coklat adalah cluster 2, dan warna abu adalah cluster 3. Pada hasil silhouette memiliki nilai rata-rata sebesar 0,738. Setelah melakukan evaluasi silhouette dengan ketentuan 4 cluster, hasil menunjukkan bahwa cluster 1 menjadi kelompok yang paling banyak memiliki anggota daripada cluster lainnya sehingga ia menjadi cluster dengan jumlah anggota paling banyak. Pada evaluasi silhouette dengan cluster 4 ini, kelompok yang memiliki jumlah anggota paling sedikit adalah cluster 3.

Tabel 3. Nilai centroid

Cluster	(V1)	(V2)	(V3)	(V4)	(V5)
Tidak Nyaman	0,525	0,5437	0,8	0,365	0,465
Nyaman	0,718	0,802	0,624	0,465	0,689
Sangat Nyaman	0,458	0,61	0,352	0,413	0,718

Tabel 3 menunjukkan nilai centroid tiap cluster di masing-masing variabel. Berdasarkan tabel di atas (V1) adalah Crime Index, (V2) adalah Safety Index, (V3) adalah GDP, (V4) adalah Air Quality, dan (V5) adalah Education Quality. Cluster 0 pada tabel menjelaskan bahwa negara pada cluster tersebut memiliki tingkat kenyamanan hidup yang rendah,

Cluster 1 pada tabel menjelaskan bahwa negara pada cluster tersebut memiliki tingkat kenyamanan hidup yang sedang, Cluster 2 pada tabel menjelaskan bahwa negara pada cluster tersebut memiliki tingkat kenyamanan hidup yang tinggi.



Gambar 5. Peta negara Asia berdasarkan tiap Cluster

Berdasarkan Gambar 5, bisa dilihat jika wilayah negara berwarna merah berarti termasuk cluster 0, jika wilayah negara berwarna kuning maka termasuk cluster 1, dan jika wilayah negara berwarna hijau berarti termasuk cluster 2. Daftar masing-masing negara bisa dilihat tabel dibawah ini.

Tabel 4. Pembagian cluster negara

Negara	Cluster
Afghanistan, Armenia, Azerbaijan, Bangladesh, Cambodia, China, Georgia, India, Indonesia, Iran, Iraq, Jordan, Kazakhstan, Kyrgyzstan, Lebanon, Malaysia, Maldives, Mongolia, Myanmar, Nepal, North Korea, Oman, Pakistan, Palestine, Philipines, Sri Lanka, Syria, Thailand, Turkey, Uzbekistan, Vietnam, Yemen	0
Qatar, Singapore	1
Bahrain, Bhutan, Brunei, Hongkong, Israel, Japan, Kuwait, Saudi Arabia, South Korea, Taiwan, United Arab Emirates	2

Tabel 4 mengelompokkan negara-negara di Asia ke dalam tiga cluster berdasarkan tingkat kenyamanan hidup yang diukur melalui indikator seperti Crime Index, Safety Index, GDP, Air Quality, dan Education Quality. Cluster 0 terdiri dari negara-negara dengan tingkat kenyamanan hidup rendah, Cluster 1 dengan tingkat kenyamanan hidup sedang, dan Cluster 2 dengan tingkat kenyamanan hidup tinggi. Hasil dari pembagian cluster ini dapat digunakan oleh pemimpin negara, pembuat kebijakan, dan peneliti untuk memahami distribusi tingkat kenyamanan hidup di Asia, serta untuk mengidentifikasi area yang memerlukan perhatian lebih lanjut dalam upaya meningkatkan kualitas hidup masyarakat. Analisis ini juga bisa menjadi dasar bagi pengembangan strategi pembangunan yang lebih efektif dan terarah di setiap negara, sesuai dengan kebutuhan dan kondisi spesifik masing-masing cluster.

### 5. KESIMPULAN

Setelah penelitian yang dilakukan menggunakan metode K-Means Clustering, peneliti mendapatkan analisis mengenai tingkat kenyamanan hidup di negara asia. Oleh karena itu, ada kesimpulan yang dapat ditarik yaitu Rentang nilai pada setiap cluster yang didapatkan adalah hasil dari implementasi metode K-Means Clustering yang ditentukan menjadi 3 cluster. Rentang nilai yang dihasilkan berasal dari pemilihan acak pada data yang digunakan dalam menghitung nilai terdekat terhadap pusat data. Maka dari itu, nilai silhouette yang dihasilkan adalah 0,761 dan hasil dari 45 data terdapat 32 data yang masuk dalam cluster rendah, 2 negara di cluster sedang, dan 11 negara di cluster tinggi.

Adapun dalam penelitian ini dilakukan ebaluasi silhouette dengan membandingkan antara 2 cluster, 3 cluster, dan 4 cluste,. Berdasarkan hasil tersebut, mengolah data dengan K-Means 3 Cluster memberikan hasil yang lebih akurat dengan Silhouette

Score 0,761. Hasil ini lebih tinggi daripada mengolah data menggunakan K-Means dengan 2 Cluster maupun 4 Cluster. Pada penelitian selanjutnya diharap dapat mengimplementasikan korelasi yang terdapat pada data menggunakan metode-metode lainnya sehingga menghasilkan kinerja lebih baik lagi.

#### DAFTAR PUSTAKA

- [1] Anton A.P. Sinaga. (2016). Analisis faktor-faktor yang mempengaruhi kesejahteraan masyarakat di kota medan (studi kasus usaha kecil dan menengah). *Jurnal Ilmiah Methonomi*, 2(1), 1–9. <http://www.methonomi.net/index.php/jimetho/article/view/6/17>
- [2] Achسانی, N. A., & Siregar, H. (2010). Classification of the ASEAN+3 economies using fuzzy clustering approach. *European Journal of Scientific Research*, 39(4), 489–497. <http://repository.ipb.ac.id/handle/123456789/30047>
- [3] Ediyanto, Mara, M. N., & Satyahadewi, N. (2013). Pengklasifikasian karakteristik dengan metode k-means cluster analysis. *Buletin Ilmiah Mat. Stat. Dan Terapannya (Bimaster)*, 02(2), 133–136. <http://dx.doi.org/10.26418/bbimst.v2i02.3033>
- [4] Dhuhita, W. M. P. (2015). Clustering Metode K-Means Untuk Menentukan Status Gizi Balita. *Jurnal Informatika*, 15(2), 160–174. <https://doi.org/10.30873/ji.v15i2.598>
- [5] Sihananto, A. N., Puspita Sari, A., Khariono, H., Akhmad Fernanda, R., & Cakra Mudra Wijaya, D. (2022). Implementasi Metode K-Means Untuk Pengelompokan Kasus Covid-19 Tingkat Provinsi Di Indonesia. *Jurnal Informatika Dan Sistem Informasi*, 3(1), 76–85. <https://doi.org/10.33005/jifosi.v3i1.472>
- [6] Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 1–8. <https://doi.org/10.36595/misi.v4i1.216>
- [7] Muchlis, A. F., & Kusuma, H. E. (2016). Persepsi kriteria kenyamanan rumah tinggal. *Temu Ilmiah IPLBI 2016*, 1, D105–D110. [https://www.researchgate.net/profile/Hanson-E-Kusuma/publication/362851245\\_Persepsi\\_Kriteria\\_Kenyamanan\\_Rumah\\_Tinggal/links/63040dafa3c7de4c34782372/Persepsi-Kriteria-Kenyamanan-Rumah-Tinggal.pdf](https://www.researchgate.net/profile/Hanson-E-Kusuma/publication/362851245_Persepsi_Kriteria_Kenyamanan_Rumah_Tinggal/links/63040dafa3c7de4c34782372/Persepsi-Kriteria-Kenyamanan-Rumah-Tinggal.pdf)
- [8] Alfiani, V. (2016). Persepsi Masyarakat Terkait Kenyamanan Tinggal di Pemukiman Kumuh (Studi Kasus: Pemukiman Kumuh Kelurahan Keputih Kecamatan Sukolilo Kota Surabaya). *Swara Bhumi*, IV(2), 8–13. <https://ejournal.unesa.ac.id/index.php/swara-bhumi/article/view/17722>
- [9] Nuryani, I., & Darwis, D. (2021). Analisis Clustering Pada Pengguna Brand Hp Menggunakan Metode K-Means. *Prosiding Seminar Nasional Ilmu Komputer*, 1(1), 190–211. <https://proceeding.unived.ac.id/index.php/snasiom/article/view/62>
- [10] Alif Fajar Fadhillah, D., Faisol, A., & Vendyansyah, N. (2022). Penerapan Metode K-Means Clustering Pada Pemetaan Lahan Kopi Di Kabupaten Malang. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(1), 162–170. <https://doi.org/10.36040/jati.v6i1.4617>
- [11] Ahmad Harmain, Paiman, P., Kurniawan, H., Kusriani, K., & Dina Maulina. (2022). Normalisasi Data Untuk Efisiensi K-Means Pada Pengelompokan Wilayah Berpotensi Kebakaran Hutan Dan Lahan Berdasarkan Sebaran Titik Panas. *TEKNIMEDIA: Teknologi Informasi Dan Multimedia*, 2(2), 83–89. <https://doi.org/10.46764/teknimedia.v2i2.49>
- [12] Mahardika, A. A., Kencana, E. N., Sukarsa, I. K. G., Jayanegara, K., Wijayakusuma, I. L., & Sumarjaya, I. W. (2023). Klasterisasi Karakteristik Wisatawan Mancanegara Menggunakan Metode K-Means Clustering. *E-Jurnal Matematika*, 12(2), 140. <https://doi.org/10.24843/mtk.2023.v12.i02.p411>
- [13] Sulistiyawati, A., & Supriyanto, E. (2021). Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan. *Jurnal Tekno Kompak*, 15(2), 25. <https://doi.org/10.33365/jtk.v15i2.1162>
- [14] Anjelita, M., Windarto, A. P., Wanto, A., & Saifullah, S. (2019). Analisis Metode K-Means pada Kasus Ekspor Barang Perhiasan dan Barang Berharga Berdasarkan Negara Tujuan. *Seminar Nasional Sains & Teknologi Informasi (SENSASI)*, 476–482. <http://prosiding.seminar-id.com/index.php/sensasi/issue/archivePage%7C476>
- [15] Metisen, B. M., & Sari, H. L. (2015). Analisis clustering menggunakan metode K-Means dalam pengelompokan penjualan produk pada Swalayan Fadhila. *Jurnal Media Infotama*, 11(2), 110–118. <https://doi.org/10.37676/jmi.v11i2.258>
- [16] Subekti, P., Andini, T. D., & Islamiyah, M. (2022). Sistem Penentuan Konsentrasi Jurusan Bagi Mahasiswa Informatika Menggunakan Metode K-Means Di Institut Asia Malang. *Jurnal Manajemen Informatika (JAMIKA)*, 12(1), 25–39. <https://doi.org/10.34010/jamika.v12i1.6452>
- [17] Orisa, M. (2022). Optimasi Cluster pada Algoritma K-Means. *Prosiding SENIATI*, 6(2), 430–437. <https://doi.org/10.36040/seniati.v6i2.5034>
- [18] Numbeo. (2024). Crime and safety index by country. Numbeo. [https://www.numbeo.com/crime/rankings\\_by\\_country.jsp?title=2024&region=142](https://www.numbeo.com/crime/rankings_by_country.jsp?title=2024&region=142)
- [19] Trading Economics. (2024). GDP per capita by country: Asia. *Trading Economics*.

- <https://tradingeconomics.com/country-list/gdp-per-capita?continent=asia>
- [20] World Population Review. (2024). Education index by country. World Population Review. <https://worldpopulationreview.com/country-rankings/education-index-by-country>
- [21] IQAir. (2024). World's most polluted countries. IQAir. <https://www.iqair.com/ca/world-most-polluted-countries>