

Pengelompokan Kabupaten/Kota berdasarkan Indikator Indeks Pembangunan Manusia 2022 Menggunakan *K-Harmonic Means Clustering*

Aya Dewanti Sofia*

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

ARTICLE INFO

Article history :

Received : 2/11/2023

Revised : 15/12/2023

Published : 25/12/2023



Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Volume : 3

No. : 2

Halaman : 163 - 172

Terbitan : Desember 2023

ABSTRAK

Clustering adalah metode untuk menemukan kluster dalam data dengan membagi data menjadi kluster yang berbeda sehingga data dari masing-masing kluster memiliki kemiripan yang maksimum sementara data antara kluster memiliki kemiripan yang minimum. Tujuan dari penelitian ini adalah mengelompokkan kabupaten/kota di Indonesia berdasarkan indikator Indeks Pembangunan Manusia (IPM) yaitu Umur Harapan Hidup (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan Pengeluaran Riil per Kapita (yang disesuaikan) pada tahun 2022 menggunakan metode klusterisasi K-Harmonic Means. Selain itu, untuk mengetahui karakteristik kabupaten/kota pada kluster yang terbentuk menggunakan metode klusterisasi K-Harmonic Means. Hasil penelitian diperoleh bahwa kluster 1 merupakan kluster dengan kategori sedang dengan anggota sebanyak 15 kabupaten/kota, kluster 2 merupakan kluster dengan kategori sangat tinggi dengan anggota sebanyak 316 kabupaten/kota, kluster 3 merupakan kluster dengan kategori tinggi dengan anggota sebanyak 182 kabupaten/kota, dan kluster 4 merupakan kluster dengan kategori rendah dengan anggota sebanyak 1 kabupaten/kota. Dari 13 kali iterasi diperoleh nilai akurasi yaitu untuk melihat kestabilan hasil klusterisasi yang terbentuk sebesar 99,4%. Dimana 511 kabupaten/kota berada pada kluster yang sama di setiap proses iterasinya, sedangkan 3 kabupaten/kota berada pada kluster yang berbeda di setiap proses iterasinya.

Kata Kunci : *Indeks Pembangunan Manusia; Klusterisasi; K-Harmonic Clustering.*

ABSTRACT

Clustering is a method to find clusters in the data by dividing the data into different clusters so that the data from each cluster has a maximum similarity while the data between clusters has a minimum similarity. The purpose of this study is to classify districts/cities in Indonesia based on Human Development Index (HDI) indicators, namely life expectancy (UHH), average school duration (RLS), School duration expectations (HLS), and real per capita expenditure (adjusted) in 2022 using the K-Harmonic Means clustering method. In addition, to determine the characteristics of districts/cities in the cluster formed using the K-Harmonic Means clustering method. The results obtained that cluster 1 is a cluster with a medium category with members as many as 15 districts/cities, cluster 2 is a cluster with a very high category with members as many as 316 districts/cities, cluster 3 is a cluster with a high category with members as many as 182 districts/cities, and Cluster 4 is a cluster with a low category with members as many as 1 district/city. From 13 iterations, the accuracy value is obtained to see the stability of the cluster formed by 99.4%. Where 511 regencies/cities are in the same cluster in each iteration process, while 3 regencies/cities are in different clusters in each iteration process.

Keywords : *Human Development Index; Clustering; K-Harmonic Clustering.*

@ 2023 Jurnal Riset Statistika, Unisba Press. All rights reserved.

A. Pendahuluan

Clustering atau klusterisasi adalah metode untuk menemukan klaster dalam data dengan membagi data menjadi klaster yang berbeda sehingga data dari setiap klaster memiliki kemiripan yang maksimum sementara data antara klaster memiliki kemiripan yang minimum. Klusterisasi adalah salah satu metode data mining dengan pembelajaran mesin yang bersifat *unsupervised learning* dimana penemuan klaster dilakukan secara otomatis menggunakan algoritma klusterisasi [1]. *K-Means* menghasilkan klaster berdasarkan posisi inialisasi klaster awal sehingga tidak menjamin hasil klusterisasi yang unik [2]. Pada *K-Harmonic Means* klaster-klaster terbentuk dari penyempurnaan iteratif yang dilihat dari letak titik pusat baru yaitu rata-rata harmonik setiap klaster.

Menurut Siagian pembangunan nasional sebagai suatu rangkaian usaha pertumbuhan dan perubahan strategis yang dilakukan oleh suatu bangsa, negara, dan pemerintah untuk tujuan kemodernan dalam skema pembinaan bangsa (*nation building*) [3] [4] [5].

Menurut *United National Development Programme*, pembangunan manusia adalah proses memperbesar kemungkinan preferensi hidup yang dimiliki manusia [6]. Indeks Pembangunan Manusia (IPM) adalah suatu upaya untuk mengukur kemajuan atau keberhasilan negara atau wilayah dalam bidang pembangunan manusia dengan menggambarkan kemampuan atau kemudahan penduduk dalam mendapatkan pendapatan, kesehatan, pendidikan, dan lainnya [7] [8]. Menurut UNDP (1990) terdapat tiga dimensi pengukur IPM yaitu umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak. Menurut BPS dengan mengacu pada pedoman UNDP, dimensi umur panjang dan hidup sehat diukur dengan umur harapan hidup, pengetahuan diukur dengan rata-rata lama sekolah dan harapan lama sekolah, sedangkan standar hidup layak diukur dengan rata-rata pengeluaran riil per kapita (yang disesuaikan). IPM Indonesia tahun 2022 yang dirilis oleh BPS pada 15 November 2022 mencapai 72,91, bertambah 0,62 poin (0,86 persen) dibandingkan tahun sebelumnya (72,29). IPM 2022 bertambah pada semua dimensi pembentuknya.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: Bagaimana hasil klusterisasi kabupaten/kota di Indonesia menggunakan metode klusterisasi *K-Harmonic Means* berdasarkan indikator Indeks Pembangunan Manusia (IPM) yaitu Umur Harapan Hidup (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan Pengeluaran per Kapita (yang disesuaikan) pada tahun 2022; Bagaimana karakteristik kabupaten/kota pada klaster yang terbentuk menggunakan metode klusterisasi *K-Harmonic Means*. Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb. (a) Mengelompokkan kabupaten/kota di Indonesia menggunakan metode klusterisasi *K-Harmonic Means* berdasarkan indikator Indeks Pembangunan Manusia (IPM) yaitu Umur Harapan Hidup (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan Pengeluaran per Kapita (yang disesuaikan) pada tahun 2022. (b) Mengetahui karakteristik kabupaten/kota pada klaster yang terbentuk menggunakan metode klusterisasi *K-Harmonic Means*.

B. Metode Penelitian

Metode Pengumpulan Data

Data pada penelitian ini adalah data sekunder mengenai Indeks Pembangunan Manusia (IPM) menurut kabupaten/kota di Indonesia yang diolah dan disajikan oleh BPS melalui situs <https://bps.go.id/>. Data tersebut adalah indikator pembentuk IPM dengan perhitungan metode baru yaitu: Umur Harapan Hidup (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan pengeluaran per kapita disesuaikan. Data yang digunakan adalah data di tahun 2022. Data yang digunakan sebanyak 514 kabupaten/kota dari 34 provinsi di Indonesia yang terdapat pada tabel 1.

Tabel 1. Data Indikator IPM Kabupaten/Kota di Indonesia Tahun 2022

No	Kabupaten/Kota	UHH (tahun)	RLS (tahun)	HLS (tahun)	Pengeluaran per Kapita (ribu rupiah /orang/tahun)
1	Simeulue	65,48	9,73	14,08	7371
2	Aceh Singkil	67,65	8,69	14,34	8994
⋮	⋮	⋮	⋮	⋮	⋮
513	Deiyai	65,66	3,26	9,84	4808
514	Kota Jayapura	70,76	11,74	15,04	15189

Metode Analisis Data

Algoritma *K-Harmonic Means (KHM) Clustering*

K-Harmonic Means (KHM) Clustering merupakan metode *clustering* berbasis pusat yang pertama kali diperkenalkan oleh Zhang, Hsu, dan Dayal pada tahun 1999 dari HP Laboratories Palo Alto yang kemudian pada tahun 2002 dikembangkan oleh Hammerly dan Elkan [9]. Metode KHM merupakan pengembangan dari metode *K-Means (KM) Clustering*, dimana keduanya membentuk kluster-kluster dengan penyempurnaan secara iteratif berdasarkan letak titik pusat (*centroid*) dari masing-masing kluster.

Secara prinsip, KHM menggunakan jumlah semua titik data dari rata-rata harmonik jarak dari titik data ke semua *centroid* sebagai fungsi obyektifnya. Rata-rata harmonik responsif dengan keadaan dimana dua atau lebih titik pusat yang saling berdekatan. Metode ini secara otomatis menempatkan satu atau lebih titik pusat ke area titik data yang jauh dari titik-titik pusat yang ada sebelumnya. Hal ini akan membuat fungsi tujuan akan semakin kecil.

Algoritma klusterisasi *K-Harmonic Means* sebagai berikut: (a) Inisialisasi posisi *centroid* awal sebanyak k *centroid* yang dihimpun dalam himpunan $C = \{c_j | j = 1, 2, \dots, k\}$ secara acak dari setiap variabel. (b) Hitung jarak setiap titik i (sebanyak n) terhadap masing-masing *centroid* dan alokasikan setiap titik i dengan melihat jarak terdekat terhadap masing-masing *centroid* ke dalam kluster sesuai dengan *centroid* tersebut. Sehingga diperoleh k buah kluster titik sebagai berikut: Kluster 1 adalah titik-titik yang jaraknya paling dekat ke *centroid* ke-1, Kluster 2 adalah titik-titik yang jaraknya paling dekat ke *centroid* ke-2, dan seterusnya, Kluster k adalah titik-titik yang jaraknya paling dekat ke *centroid* ke- k . (c) Misal D adalah himpunan obyek dalam kluster ke- j dengan $j = 1, 2, \dots, k$, dan $|D_j|$ adalah banyaknya anggota kluster ke- j . Hitung *centroid* baru dari setiap kluster dengan menggunakan rata-rata harmonik setiap variabel X sebanyak p , sebagai berikut:

Rata-rata harmonik variabel X_1 di kluster ke- j

$$HM_{X_1, d_j} = \frac{|D_j|}{\sum_{i \in D_j} \frac{1}{x_{i1}}} \tag{1}$$

Rata-rata harmonik variabel X_p di kluster ke- j

$$HM_{X_p, d_j} = \frac{|D_j|}{\sum_{i \in D_j} \frac{1}{x_{ip}}} \tag{2}$$

Misal himpunan $E = \{e_j | j = 1, 2, \dots, k\}$ adalah kumpulan *centroid* baru sebanyak k *centroid* yang diambil dari hasil iterasi pada langkah 3 dimana akan digunakan untuk iterasi selanjutnya.

Ulangi langkah 3 sampai 5 hingga posisi kluster sudah stabil dimana obyek i tidak mengalami perubahan alokasi kluster dari iterasi ke iterasi selanjutnya. Kriteria kestabilan ini dilihat melalui *confusion matrix* menggunakan nilai akurasi dari rentang 0 sampai 1 dimana semakin mendekati nilai 1 maka kestabilan kluster semakin baik.

C. Hasil dan Pembahasan

Analisis Deskriptif

Analisis deskriptif dilakukan dengan tujuan untuk mendeskripsikan atau memberikan gambaran umum dari data yang diteliti tanpa melakukan analisis dan membuat kesimpulan [10] [11].

Tabel 2. Statistika Deskriptif Indikator IPM

Variabel	Minimum	Median	Mean	Maksimum
Umur Harapan Hidup (X_1)	55,7	70,22	69,93	77,82
Rata-rata Lama Sekolah (X_2)	1,580	8,410	8,552	13,030
Harapan Lama Sekolah (X_3)	4,07	13,01	13,09	17,81
Pengeluaran per kapita (X_4)	4190	10512	10643	24221

Kabupaten/kota dengan Umur Harapan Hidup, Rata-rata Lama Sekolah, Harapan Lama Sekolah, Pengeluaran per kapita terendah adalah kabupaten Nduga. Sedangkan kabupaten/kota dengan Umur Harapan Hidup tertinggi adalah kabupaten Sukoharjo, kabupaten/kota dengan Rata-rata Lama Sekolah tertinggi adalah kabupaten Bulungan, kabupaten/kota dengan Harapan Lama Sekolah tertinggi adalah kota Banda Aceh, dan kabupaten/kota dengan Pengeluaran per kapita tertinggi adalah kota Jakarta Selatan.

Nilai median dan *mean* setiap variabel cukup berbeda jauh dengan nilai minimum/maksimum pada setiap variabel indikator IPM. Hal tersebut dapat menjadi indikasi bahwa kabupaten/kota di Indonesia mengalami kesenjangan pada setiap indikator IPM nya.

Persiapan Data

Data yang digunakan dalam penelitian memiliki empat variabel dengan skala pengukuran yang berbeda, maka perlu dilakukan standarisasi skala pengukuran menggunakan normalisasi *Min-max* pada keempat variabel indikator IPM sehingga ukuran yang digunakan tidak berbeda dan memenuhi standar normal dimana nilai nya diantara 0 sampai dengan 1 [12].

Tabel 3. Data Normalisasi Min-max

No	Kabupaten/Kota	Data Terstandarisasi			
		X_1	X_2	X_3	X_4
1	Simeulue	0,4421338	0,7117904	0,7285298	0,15880385
2	Aceh Singkil	0,5402351	0,6209607	0,7474527	0,23982827
⋮	⋮	⋮	⋮	⋮	⋮
513	Deiyai	0,4502712	0,14672489	0,4199418	0,03085218
514	Kota Jayapura	0,6808318	0,88733624	0,7983988	0,54909890

Algoritma K-Means Harmonic Clustering

Inisialisasi Posisi Centroid Awal

Dalam menentukan posisi *centroid* awal dimana kluster yang akan dibentuk sebanyak empat kluster ($k = 4$), maka posisi *centroid* yang akan ditentukan sebanyak empat *centroid* untuk empat kluster pada empat variabel indikator IPM.

Tabel 4. Kluster 1 Hasil Pengacakan

No	Data ke-	X_1	X_2	X_3	X_4
1	129	0,6921338	0,5641921	0,6470160	0,31016924
2	509	0,4529837	0,1406114	0,3951965	0,03055264
⋮	⋮	⋮	⋮	⋮	⋮
129	413	0,6532550	0,5641921	0,6681223	0,39498777

Data kluster 2 dimana sebanyak 129 data terdapat pada Tabel 4.

Tabel 5. Kluster 2 Hasil Pengacakan

No	Data ke-	X_1	X_2	X_3	X_4
1	411	0,5479204	0,4873362	0,6717613	0,42628925
2	163	0,6726944	0,4908297	0,5778748	0,20238630
⋮	⋮	⋮	⋮	⋮	⋮
129	407	0,5162749	0,5650655	0,6339156	0,38075982

Data klaster 3 dimana sebanyak 128 data terdapat pada Tabel 5.

Tabel 6. Klaster 3 Hasil Pengacakan

No	Data ke-	X_1	X_2	X_3	X_4
1	399	0,5926763	0,56855895	0,6259098	0,26239329
2	352	0,6781193	0,58253275	0,6128093	0,40072887
⋮	⋮	⋮	⋮	⋮	⋮
128	407	0,8955696	0,7912664	0,7248908	0,66172433

Data klaster 4 dimana sebanyak 128 data terdapat pada Tabel 6.

Tabel 7. Klaster 4 Hasil Pengacakan

No	Data ke-	X_1	X_2	X_3	X_4
1	448	0,7007233	0,5606987	0,6157205	0,1820678
2	420	0,7549729	0,8707424	0,8384279	0,6597773
⋮	⋮	⋮	⋮	⋮	⋮
128	54	0,7726040	0,8384279	0,7132460	0,35784534

Selanjutnya data setiap variabel indikator IPM pada setiap klaster tersebut dipilih secara acak untuk dijadikan *centroid* awal klaster.

Tabel 8. Centroid Awal pada Setiap Klaster

Klaster	Variabel			
	X_1	X_2	X_3	X_4
Klaster 1	0,5583183	0,2951965	0,7234352	0,2396785
Klaster 2	0,7400542	0,5065502	0,6754003	0,7146922
Klaster 3	0,8318264	0,6384279	0,6135371	0,2348859
Klaster 4	0,5754973	0,1467249	0,7132460	0,2230043

Menghitung Jarak dan Menentukan Jarak Terdekat ke setiap *Centroid*

Langkah selanjutnya adalah menghitung jarak setiap titik i pada setiap variabel indikator IPM yaitu sebanyak 514 titik terhadap masing-masing *centroid* awal yang sudah ditentukan secara acak menggunakan pengukuran jarak *Euclidean*.

Tabel 9. Hasil Jarak Euclidean

Data ke-	Klaster 1	Klaster 2	Klaster 3	Klaster 4	Klaster Terdekat
1	0,4400181	0,6653681	0,4198264	0,5843288	3
2	0,3271485	0,5326395	0,3213851	0,4770704	3
⋮	⋮	⋮	⋮	⋮	⋮
514	0,6833737	0,4370985	0,4665387	0,8204811	2

Berdasarkan hasil perhitungan jarak *Euclidean* diatas, diperoleh jarak minimum atau jarak terdekat terhadap masing-masing *centroid* klaster yang menandakan keanggotaan klaster baru yang akan digunakan untuk menghitung *centroid* baru untuk iterasi selanjutnya menggunakan rata-rata harmonik.

Tabel 10. Jumlah Anggota Setiap Klaster

Klaster	Jumlah Anggota
1	138
2	47
3	318
4	11

Centroid Baru dengan Harmonic Means

Keanggotan hasil klasterisasi selanjutnya digunakan untuk menghitung *centroid* baru menggunakan rata-rata harmonik

Tabel 11. Centroid Baru pada Setiap Klaster

Klaster	Variabel			
	X ₁	X ₂	X ₃	X ₄
Klaster 1	0,4413368	0,5021706	0,6273698	0,1852265
Klaster 2	0,785424	0,7785774	0,7440416	0,5911561
Klaster 3	0,6809750	0,6214749	0,6589346	0,3002116
Klaster 4	0	0	0	0

Sehingga dalam menentukan keanggotaan klaster baru, jarak terdekat diukur terhadap masing-masing *centroid* yang diperoleh dari rata-rata harmonik.

Tabel 12. Hasil Jarak Euclidean Iterasi 1

Data ke-	Klaster 1	Klaster 2	Klaster 3	Klaster 4	Klaster Terdekat
1	0,23424904	0,55630736	0,30006940	1,1216516	1
2	0,20320731	0,45651244	0,17688864	1,1373870	3
⋮	⋮	⋮	⋮	⋮	⋮
514	0,60610601	0,1658058	0,38997111	1,4798135	2

Berdasarkan hasil perhitungan jarak *Euclidean* diatas terlihat bahwa klaster baru yang terbentuk mengalami perubahan keanggotan klaster dari iterasi sebelumnya setelah menggunakan rata-rata harmonik sebagai penentuan *centroid* baru.

Tabel 13. Jumlah Anggota Setiap Klaster pada Iterasi 1

Klaster	Jumlah Anggota
1	129
2	84
3	298
4	3

Dikarenakan keanggotaan kluster masih mengalami perubahan yang signifikan maka perlu dilakukan iterasi kembali hingga tidak ada perubahan alokasi kluster.

Pengulangan Iterasi dan Pengelompokan Kluster

Pengulangan iterasi dilakukan 13 kali dimana hasil alokasi anggota kluster disajikan pada Tabel 14.

Tabel 14. Alokasi Kluster pada Proses Iterasi Kluster

Data ke-	Inisialisasi Awal	Iterasi ke-												
		1	2	3	4	5	6	7	8	9	10	11	12	13
1	3	1	1	3	3	3	3	3	3	3	3	3	3	3
2	3	3	3	3	3	3	3	3	3	3	3	3	3	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
514	2	2	2	2	2	2	2	2	2	2	2	2	2	2

Berdasarkan tabel diatas dapat dilihat bahwa posisi kluster cenderung sudah stabil atau tidak mengalami perubahan anggota kluster. Sehingga dari 13 kali iterasi, keanggotaan setiap variabel indikator IPM dibagi ke dalam empat kluster menurut Tabel 15.

Tabel 15 Alokasi Anggota Hasil Klasterisasi

Kluster	Jumlah Anggota
1	15
2	316
3	182
4	1

Kestabilan Kluster

Dalam melihat kestabilan kluster yang sudah dibentuk melalui 13 kali iterasi, dibutuhkan suatu ukuran untuk menilai seberapa stabil posisi kluster pada setiap iterasi.

Tabel 16. Nilai Akurasi setiap Iterasi

Iterasi ke-	Nilai Akurasi	Iterasi ke-	Nilai Akurasi
1	0,8210117	8	0,9416342
2	0,9066148	9	0,9396887
3	0,9260700	10	0,9435798
4	0,9221790	11	0,9669261
5	0,9299611	12	0,9708171
6	0,9357977	13	0,9941634
7	0,9377432		

Dengan angka akurasi yang ditentukan yaitu sebesar 0,99, dapat dilihat bahwa nilai akurasi dapat mencapai angka akurasi 0,99 dan mendekati angka 1 seiring dengan proses iterasi, yang artinya kluster yang dibentuk semakin stabil atau cenderung tidak ada perubahan posisi kluster. Terlihat pada tabel bahwa pada

iterasi ke-13 nilai akurasi bernilai 0,9941634 atau 99,4% artinya sekitar 511 data (511 kabupaten/kota) indikator IPM berada di klaster yang sama pada setiap iterasinya sementara sisanya yaitu 0,6% atau sekitar 3 data (3 kabupaten/kota) indikator IPM berada di klaster yang berbeda pada setiap iterasinya.

Identifikasi Karakteristik Setiap Klaster

Identifikasi karakteristik merupakan proses interpretasi hasil klasterisasi dengan memberi label atau nama pada setiap klaster yang terbentuk yang dapat menggambarkan sifat dari setiap klaster.

Tabel 17. Rata-rata Hasil Klasterisasi

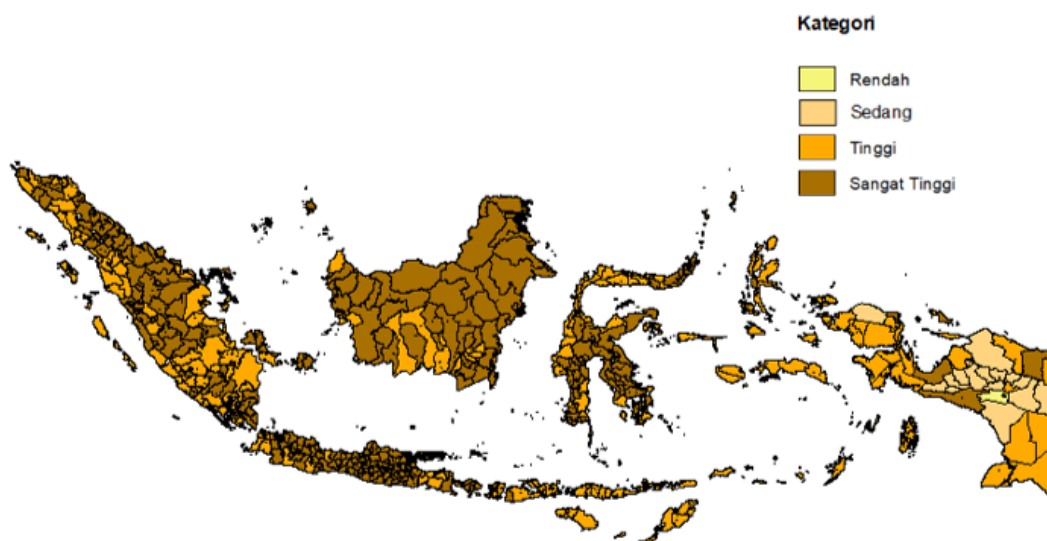
Centroid Klaster	Variabel				Kategori
	Umur Harapan Hidup	Rata-rata Lama Sekolah	Harapan Lama Sekolah	Pengeluaran per Kapita	
1	64,5053	4,0680	9,0967	5282,3333	Sedang
2	71,9564	9,1536	13,4243	11898,7563	Sangat Tinggi
3	66,934	7,916	12,884	8940,857	Tinggi
4	55,7	1,58	4,07	4190	Rendah

Berdasarkan tabel diatas terlihat bahwa setiap klaster memiliki *centroid* yang perbedaannya cukup besar untuk menunjukkan perbedaan atau karakteristik setiap klasternya. Dari empat klaster tersebut dibagi menjadi empat kategori untuk indikator IPM setiap kabupaten/kota yaitu kategori dengan rata-rata klaster sangat tinggi, tinggi, sedang, dan rendah.

Berdasarkan hasil identifikasi karakteristik klaster yang dijelaskan, terlihat bahwa terjadi kesenjangan kesehatan, pendidikan, dan ekonomi pada setiap kabupaten/kota. Dari hasil klasterisasi pun terlihat bahwa jumlah anggota antar klaster jauh berbeda

Visualisasi Hasil Klaster

Visualisasi hasil klaster dilakukan untuk melihat penyebaran kabupaten/kota di Indonesia menurut karakteristik setiap klasternya. Visualisasi tersebut digambarkan menggunakan diagram peta (kartogram) pada Gambar 1.



Gambar 1. Diagram Peta Hasil Klasterisasi

Berdasarkan diagram peta diatas dapat terlihat pulau mana yang harus menjadi perhatian pemerintah untuk meningkatkan dan mengembangkan setiap indikator pembangun IPM agar tidak terjadinya kesenjangan antar wilayah kabupaten/kota di pulau atau pun provinsi tersebut menurut karakteristik daerah nya masing-masing.

D. Kesimpulan

Berdasarkan identifikasi masalah pada penelitian ini, dapat disimpulkan bahwa metode *K-Harmonic Means Clustering* telah mengelompokkan 514 kabupaten/kota di Indonesia menurut indikator IPM pada tahun 2022 menjadi empat klaster dimana karakteristik setiap klaster dikategorikan menjadi empat kategori menurut rata-rata (*centroid*) setiap klasternya, yaitu rendah, sedang, tinggi, dan sangat tinggi. (1) Klaster 1 merupakan klaster dengan nilai rata-rata per variabelnya lebih kecil dari klaster 2 dan 3 tetapi lebih besar dari klaster 4. Sehingga klaster 1 merupakan klaster dengan rata-rata (*centroid*) sedang. Terdapat 15 kabupaten/kota yang termasuk dalam klaster 1, artinya terdapat 15 kabupaten/kota di Indonesia yang umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran per kapita nya sedang di tahun 2022. (2) Klaster 2 merupakan klaster dengan nilai rata-rata per variabelnya lebih besar dari klaster 1, 3, dan 4 atau klaster dengan rata-rata per variabelnya tertinggi dari klaster lainnya. Sehingga klaster 2 merupakan klaster dengan rata-rata (*centroid*) sangat tinggi. Terdapat 316 kabupaten/kota yang termasuk dalam klaster 2, artinya terdapat 316 kabupaten/kota di Indonesia yang umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran per kapita nya sangat tinggi di tahun 2022. (4) Klaster 3 merupakan klaster dengan nilai rata-rata per variabelnya lebih kecil dari klaster 2 tetapi lebih besar dari klaster 1 dan 4. Sehingga klaster 3 merupakan klaster dengan rata-rata (*centroid*) tinggi. Terdapat 182 kabupaten/kota yang termasuk dalam klaster 3, artinya terdapat 182 kabupaten/kota di Indonesia yang umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran per kapita nya tinggi di tahun 2022. (5) Klaster 4 merupakan klaster dengan nilai rata-rata per variabelnya lebih kecil dari klaster 2, 3, dan 4 atau klaster dengan rata-rata per variabelnya terendah dari klaster lainnya. Sehingga klaster 4 merupakan klaster dengan rata-rata (*centroid*) rendah. Terdapat 1 kabupaten/kota yang termasuk dalam klaster 1, artinya terdapat 1 kabupaten/kota di Indonesia yang umur harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran per kapita nya rendah di tahun 2022.

Daftar Pustaka

- [1] Batta. M, "Machine Learning Algorithms - A Review," *International Journal of Science and Research (IJSR)*, vol. 18, no. 8, pp. 381–386, 2018.
- [2] S. S. Khan and A. Ahmad, "Cluster center initialization algorithm for K-means clustering," *Pattern recognition letters*, vol. 25, no. 11, pp. 1293–1302, 2004.
- [3] S. P. Siagian, *Proses Pengelolaan Pembangunan Nasional*. Jakarta: Penerbit CV." Haji Masagung, 1994.
- [4] T. W. Astuti, "Analisis Pengaruh Pengeluaran Pemerintah pada Sektor Kesejahteraan terhadap Pembangunan Manusia Kabupaten/Kota Sumatera Utara," Universitas Negeri Medan, Medan, 2013.
- [5] A. Faizal, "Motivasi Belajar Intrinsik dan Ekstrinsik Siswa Jurusan Teknik Kendaraan Ringan di SMKN 1 Cangkringan," Universitas Negeri Yogyakarta, Yogyakarta, 2019.
- [6] S. Sukirno, *Pengantar Teori Makro Ekonomi*. Jakarta: Raja Grafindo Persada, 2006.
- [7] BPS, "Indeks Pembangunan Manusia," 2022.
- [8] D. Daria, "Analisis Kelompok Wilayah Rawan Penyakit Malaria di Provinsi Nusa Tenggara Timur Tahun 2014," Universitas Islam Indonesia, Yogyakarta, 2016.
- [9] Zhang. B, Hsu. M, and Dayal. U, "K-Harmonic means - A data clustering algorithm," *HP Laboratories Technical Report*, p. 124, 1999.
- [10] Sugiyono, *Metode Penelitian Kuantitatif Kualitatif*. Penerbit Alfabeta cv, 2019.

- [11] E. F. Dewi and N. Hajarisman, “Penanganan Data Hilang pada Pemodelan Persamaan Terstruktur melalui Metode Full Information Maximum Likelihood (FIML),” *DataMath: Journal of Statistics and Mathematics*, vol. 1, no. 1, pp. 11–18, 2023, doi: 10.29313/datamath.v1i1.10.
- [12] A. Firdaus, “Aplikasi Algoritma K-Nearest Neighbor pada Analisis Sentimen Omicron Covid-19,” *Jurnal Riset Statistika*, pp. 85–92, Dec. 2022, doi: 10.29313/jrs.v2i2.1148.