

# Penentuan Klustering Indeks Pembangunan Manusia Provinsi Jawa Tengah dengan Metode K-Means Berbasis Web

Rakryan Aryasatya <sup>1\*</sup>, Veronica Lusiana <sup>2</sup>

<sup>1,2</sup> Program Studi Teknik Informatika, Fakultas Teknologi Informasi dan Industri, Universitas Stikubank, Kota Semarang, Provinsi Jawa Tengah, Indonesia.

## article info

### Article history:

Received 18 July 2023

Received in revised form

6 November 2023

Accepted 25 November 2023

Available online January 2024

### DOI:

<https://doi.org/10.35870/jti.k.v8i1.1403>

### Keywords:

System; Human Development Index; K-Means.

### Kata Kunci:

Sistem; Indeks Pembangunan Manusia; K-Means.

## abstract

The Human Development Index uses a data clustering algorithm, namely the K-Means algorithm, which is the simplest clustering algorithm compared to other algorithms. This algorithm is one of the most important algorithms in data mining. K-Means divides the data and then groups it into several similar clusters and separates each cluster based on the differences between each cluster. The aim of this research is to design and implement the Human Development Index for Central Java Province using a web-based k-means clustering algorithm. This research is a qualitative research in the field of electrical engineering, especially in the field of software. This research was conducted by analyzing data using the K-Means Clustering Algorithm for the Human Development Index. The implementation of the k-means clustering algorithm into the clustering system provides effective data grouping classification results and the process of each centroid distance rotation literacy, the determination of cluster points is formed, human data as a reference object saves more time when clustering the Human Development Index. The application of this clustering results in more flexible information that can be accessed at any time by users who are given access rights to utilize the data. The application of the K-Means Clustering Algorithm to obtain the results of the Human Development Index requires an information system implementation to form four clusters.

## abstract

Indeks Pembangunan Manusia yang menggunakan algoritma data clustering yaitu algoritma K-Means merupakan algoritma clustering yang paling sederhana dibanding dengan algoritma yang lain. Algoritma ini termasuk salah satu algoritma paling penting dalam data mining. K-Means membagi data kemudian mengelompokkannya ke dalam beberapa cluster yang memiliki kemiripan dan memisahkan setiap cluster berdasarkan perbedaan antar masing-masing cluster. Tujuan dalam penelitian ini adalah merancang dan menerapkan Indeks Pembangunan Manusia Provinsi Jawa Tengah menggunakan algoritma klastering k-means berbasis web. Penelitian ini merupakan penelitian kualitatif dalam bidang teknik elektro khususnya bidang perangkat lunak. Penelitian ini dilakukan dengan cara menganalisis data dengan menggunakan Algoritma Clustering K-Means Indeks Pembangunan Manusia. Implementasi algoritma k-means clustering ke dalam sistem klasterisasi memberikan hasil klasifikasi pengelompokan data yang efektif dan proses setiap literasi perputaran jarak centroid, penentuan titik cluster dibentuk, data manusia sebagai acuan objek lebih menghemat waktu melakukan klasterisasi Indeks Pembangunan Manusia. Penerapan klasterisasi ini menghasilkan informasi yang lebih fleksibel dapat diakses kapan saja oleh pengguna yang diberi hak akses untuk memanfaatkan datanya. Penerapan Algoritma K-Means Clustering mendapatkan hasil Indeks Pembangunan Manusia dibutuhkan suatu implementasi sistem informasi terbentuk empat cluster.

\*Corresponding Author. Email: [arsatya.24@gmail.com](mailto:arsatya.24@gmail.com) <sup>1\*</sup>.

© E-ISSN: 2580-1643.

Copyright © 2024 by the authors of this article. Published by Lembaga Otonom Lembaga Informasi dan Riset Indonesia (KITA INFO dan Riset). This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. 



ACM Computing Classification System (CCS)



Communication and Mass Media Complete (CMMC)

## 1. Latar Belakang

Keberhasilan suatu negara dalam pembangunan nasional tidak hanya dilihat dari pertumbuhan ekonomi akan tetapi juga dilihat dari peningkatan kualitas sumber daya manusia dalam pembangunan manusia. Keberhasilan dalam upaya membangun kualitas hidup manusia dapat diukur dengan Indeks Pembangunan Manusia (IPM). IPM menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam bentuk pendapatan, kesehatan, pendidikan, dan aspek lain dalam kehidupan. Indeks Pembangunan Manusia (IPM) pertama kali diperkenalkan oleh *United Nations Development Program* (UNDP) pada tahun 1990 dan di publikasikan dalam laporan tahunan *Human Development Report* atau (HDR).

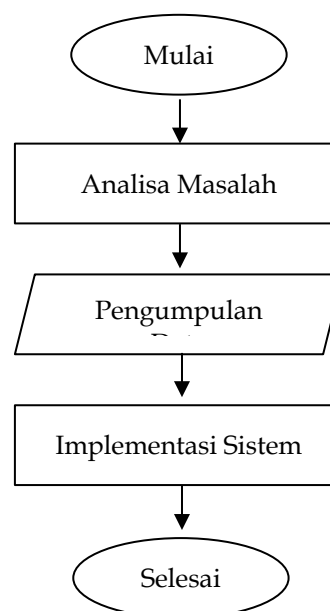
Analisis kluster merupakan salah satu teknik yang penting dalam pengolahan data yang bertujuan untuk mengelompokkan data menjadi beberapa kluster berdasarkan kesamaan atau perbedaan karakteristik tertentu. Salah satu algoritma yang umum digunakan untuk melakukan analisis kluster adalah algoritma K-Means. Penelitian-penelitian yang telah dilakukan sebelumnya di Indonesia menunjukkan beragam penerapan algoritma K-Means dalam konteks indeks pembangunan manusia (IPM) di beberapa wilayah. Penelitian yang dilakukan terhadap provinsi Sumatera Barat telah mengaplikasikan algoritma K-Means untuk menganalisis data IPM kabupaten/kota pada tahun 2021. Hasilnya menunjukkan bahwa terdapat tiga kluster yang terbentuk, yaitu kluster dengan IPM rendah, kluster dengan IPM tinggi, dan kluster dengan IPM menengah di antara keduanya. Hal ini memberikan gambaran tentang disparitas pembangunan manusia di wilayah tersebut [1].

Penelitian serupa juga dilakukan di Pulau Kalimantan, di mana algoritma K-Means digunakan untuk mengelompokkan kabupaten/kota berdasarkan IPM. Penggunaan metode ini menghasilkan empat kluster dengan tingkat keseragaman yang sedang, seperti yang diindikasikan oleh nilai koefisien Silhouette yang diperoleh [2]. Algoritma K-Means juga diterapkan untuk mengelompokkan kabupaten/kota berdasarkan IPM. Hasilnya menunjukkan adanya tiga kluster dengan tingkat IPM yang berbeda-beda, dari rendah, sedang, hingga tinggi [3]. Penelitian yang lebih luas

mencakup seluruh wilayah Indonesia menunjukkan bahwa algoritma K-Means dapat digunakan untuk mengelompokkan kabupaten/kota berdasarkan IPM dengan baik. Hasil pengelompokan menunjukkan adanya empat kluster yang mencerminkan tingkat pembangunan manusia yang berbeda-beda, mulai dari rendah hingga sangat tinggi [4]. Penerapan teknik data mining dalam analisis kluster IPM provinsi di Indonesia menunjukkan bahwa terdapat enam kluster yang terbentuk, mencakup berbagai tingkat IPM dari yang sangat baik hingga sangat buruk. Evaluasi kinerja menggunakan Davies-Bouldin Index menegaskan bahwa jumlah kluster terbaik dalam konteks ini adalah enam [5]. Dari rangkaian penelitian ini, dapat disimpulkan bahwa algoritma K-Means memiliki fleksibilitas dan efektivitas yang baik dalam mengelompokkan data IPM di berbagai wilayah di Indonesia. Hasil-hasil penelitian ditujukan kepada pengambil kebijakan dalam merumuskan strategi pembangunan yang lebih terarah dan berkelanjutan serta menjadi landasan dalam analisis kluster dan pembangunan manusia di Indonesia.

## 2. Metode Penelitian

Langkah-langkah yang dilakukan dalam perancangan/pembuatan pada penelitian ini secara garis besarnya dapat dibagi dalam berbagai tahapan yang dimulai dengan analisis, pengumpulan data, dan Implementasi Algoritma K-Means.



Gambar 1. Flowchart Penelitian

Gambar 1 merupakan Flowchart yang nanti akan dijadikan penelitian dalam membuat aplikasi Klastering Indeks Pembangunan Manusia dengan Metode K- Means Berbasis Website. Tahapan dari Flowchart ini yaitu :

#### *Analisis Masalah*

Proses clustering data indeks pembangunan manusia masih dilakukan dengan metode lama. Pengguna sulit mendapatkan informasi data indeks pembangunan manusia. apabila ingin bertanya tentang data indeks pembangunan manusia secara lengkap mereka harus datang ke kantor BPS sehingga itu cenderung lambat dan memakan waktu Sistem baru yang diharapkan sesuai kebutuhan dan memiliki data yang lengkap sehingga memudahkan pengguna diharapkan mampu digunakan secara maksimal. Penelitian ini merupakan penelitian kualitatif dalam bidang teknik elektro khususnya bidang perangkat lunak. Penelitian ini dilakukan dengan cara membangun suatu perangkat lunak (software) berupa penentuan *clustering* Indeks Pembangunan Manusia Provinsi Jawa Tengah menggunakan algoritma clustering k-means.

#### *Pengumpulan Data*

Adapun sumber data yang digunakan dalam penelitian ini dibagi menjadi dua yaitu :

- 1) *Data Primer* adalah data yang diperoleh secara langsung dari obyek penelitian melalui pengamatan langsung di lapangan mengenai hal-hal yang berkaitan dengan indikator penelitian. Data primer penelitian ini dikumpulkan melalui metode yaitu pengumpulan data yang berhubungan dengan penentuan clustering indeks pembangunan manusia Provinsi Jawa Tengah
- 2) *Data Sekunder* adalah data pendukung yang diperoleh dari berbagai sumber. Data sekunder penelitian ini dikumpulkan melalui studi literatur, yaitu penelusuran literatur mengenai dasar pengetahuan tentang hal-hal yang berkaitan dengan penelitian ini. Metode ini dilakukan dengan cara mencari buku-buku, artikel-artikel, dan jurnal-jurnal ilmiah mengenai sistem informasi, rekayasa perangkat lunak, pemrograman web dan bahasa pemrograman.

#### *Tahapan Pengembangan Sistem*

Tahapan pengembangan sistem Penentuan Clustering Indeks Pembangunan Manusia dengan Metode K-Means antara lain:

- 1) Mengumpulkan data indeks pembangunan manusia dari sumber yang terpercaya dan sesuai dengan kebutuhan serta memastikan data tersebut bersifat lengkap dan terbaru.
- 2) Mempersiapkan data dengan menghapus data yang tidak lengkap atau duplikat, menormalisasi data jika diperlukan, dan melakukan seleksi fitur untuk menentukan variabel apa saja yang akan digunakan dalam analisis.
- 3) Menentukan jumlah klaster yang akan dibentuk dengan metode K-Means. Pemilihan k dapat dilakukan dengan menggunakan Pemilihan k dapat dilakukan dengan menggunakan metode *silhouette*, yaitu dengan menentukan garis tertinggi atau melihatnya dengan melihat garis yang paling optimum.
- 4) Menentukan lokasi awal centroid secara acak di dalam wilayah data.
- 5) Menghitung jarak setiap data point terhadap setiap centroid.
- 6) Menentukan klaster yang paling dekat dengan setiap data point, dan memasukkan data point ke dalam klaster tersebut.
- 7) Menghitung ulang posisi *centroid* untuk setiap klaster, berdasarkan rata-rata posisi data point di dalam kluster sampai posisi centroid tidak berubah atau telah mencapai batas iterasi yang telah ditentukan.
- 8) Mengevaluasi hasil clustering dengan menggunakan metode yang sesuai, misalnya dengan menghitung SSE atau menggunakan metode *silhouette*.

#### *Proses Implementasi Algoritma K-Means*

Algoritma k-means pada dasarnya melakukan tiga proses yaitu, proses pendeteksian lokasi pusat cluster, proses pencarian anggota dari tiap-tiap kluster dan proses pencarian anggota dari tiap-tiap cluster. Proses Algoritma k-means sebagai berikut :

- 1) Penentuan pusat *cluster* awal  
Dalam menentukan n buah pusat cluster awal dilakukan pembangkitan bilangan random yang merepresentasikan urutan data input. Pusat awal cluster didapatkan dari data nilai max dan nilai min bukan dengan menentukan titik baru secara dengan random pusat awal dari data.

2) Perhitungan jarak dengan pusat *cluster*

Untuk mengukur jarak antar data dengan pusat *cluster* digunakan algoritma euclidian distance, algoritma perhitungan jarak data dengan pusat *cluster* :

- Ambil nilai data dan nilai pusat *cluster*.
- Hitung city block distance dengan data tiap pusat *cluster* menggunakan persamaan (1).

$$d_{ij} = \sqrt{\sum_{k=1}^n |x_{ik} - x_{jk}|} \quad (1)$$

Dimana  $X_{ik}$  adalah pusat *cluster* baru,  $X_{jk}$  pada kasus ini adalah nilai dari penduduk  $d_{ij}$  merupakan jarak,  $i$  merupakan pusat *cluster* di mana  $i = 1, 2, \dots, n$ ,  $j$  merupakan data nilai dari setiap kota / kabupaten, di mana  $j = 1, 2, \dots, n$ .  $n$  merupakan jumlah sampel.

3) Pengelompokan data

Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat. Algoritma pengelompokan data :

- Ambil nilai jarak tiap pusat *cluster* dengan data.
- Cari nilai jarak terkecil.
- Kelompokkan data dengan pusat *cluster* yang memiliki jarak yang terkecil.

4) Penentuan pusat *cluster* baru

Untuk mendapatkan pusat *cluster* baru bisa dihitung dari rata-rata nilai anggota *cluster* dan pusat *cluster*. Pusat *cluster* yang baru digunakan untuk melakukan iterasi selanjutnya, jika hasil yang didapatkan belum konvergen. Proses iterasi akan berhenti jika telah memenuhi maksimum iterasi yang dimasukkan oleh user atau hasil yang dicapai sudah konvergen (pusat *cluster* baru sama dengan pusat *cluster* lama). Algoritma penentuan pusat *cluster* :

- Cari jumlah anggota tiap *cluster*.
- Hitung pusat *cluster* baru menggunakan persamaan (2) :

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j \quad (2)$$

Dimana  $1/M$  merupakan jumlah data dari setiap *cluster*,  $C_i$  merupakan fitur ke- $i$  dalam sebuah *cluster*, dan  $X$  merupakan data dari setiap *cluster*.

5) Ulangi langkah ketiga, jika *cluster* masih berubah Karakteristik algoritma k-means berdasarkan cara kerjanya sebagai berikut :

- K-Means* sangat cepat dalam proses *clustering*.
- K-Means* sangat sensitif pada proses pembangkitan *centroid* awal secara random.
- Memungkinkan suatu *cluster* tidak mempunyai anggota.
- Hasil *clustering* dengan k-means bersifat tidak unik (selalu berubah), terkadang baik, dan terkadang buruk.

### 3. Hasil dan Pembahasan

Berikut disajikan data yang akan digunakan untuk mengelompokkan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan Indeks Pembangunan Manusia tahun 2022 bisa dilihat pada tabel 1.

Tabel 1. Data Indikator

Kabupaten / Kota	UHH	HLS	RLS	PRP
Kabupaten Cilacap	74,07	12,66	7,18	10904,00
Kabupaten Banyumas	73,88	13,21	7,78	11905,00
Kabupaten Purbalingga	73,28	12,01	7,33	10277,00
Kabupaten Banjarnegara	74,37	11,81	6,84	9776,00
Kabupaten Kebumen	73,70	13,36	7,85	9282,00
Kabupaten Purworejo	75,03	13,52	8,32	10671,00
Kabupaten Wonosobo	75,05	11,78	6,88	11108,00
Kabupaten Magelang	74,03	12,58	7,81	10011,00
Kabupaten			8,08	13250,00

Boyolali	76,12	12,62		
Kabupaten Klaten	76,95	13,40	9,09	12522,00
Kabupaten Sukoharjo	77,82	13,90	9,62	11841,00
Kabupaten Wonogiri	76,41	12,51	7,42	9780,00
Kabupaten Karanganyar	77,64	13,70	8,79	11798,00
Kabupaten Sragen	75,87	12,91	7,79	13052,00
Kabupaten Grobogan	74,93	12,45	7,26	10610,00
Kabupaten Blora	74,60	12,44	7,01	10067,00
Kabupaten Rembang	74,68	12,13	7,41	10937,00
Kabupaten Pati	76,32	12,95	7,79	10948,00
Kabupaten Kudus	76,36	13,25	9,06	11609,00
Kabupaten Jepara	75,97	12,77	8,09	10913,00
Kabupaten Demak	75,52	13,33	8,10	10698,00
Kabupaten Semarang	75,86	13,04	8,05	12448,00
Kabupaten Temanggung	75,70	12,55	7,41	9773,00
Kabupaten Kendal	74,53	12,97	7,71	11999,00
Kabupaten Batang	74,79	12,14	6,90	9972,00
Kabupaten Pekalongan	73,80	12,43	7,46	10707,00
Kabupaten Pemalang	73,65	11,98	6,50	8994,00
Kabupaten Tegal	71,85	12,91	7,25	10020,00
Kabupaten Brebes	69,74	12,51	6,35	10514,00
Kota Magelang	77,02	14,31	10,94	12816,00
Kota Surakarta	77,43	14,89	10,92	15463,00
Kota Salatiga	77,72	15,43	10,95	16351,00
Kota Semarang	77,69	15,54	10,80	16047,00

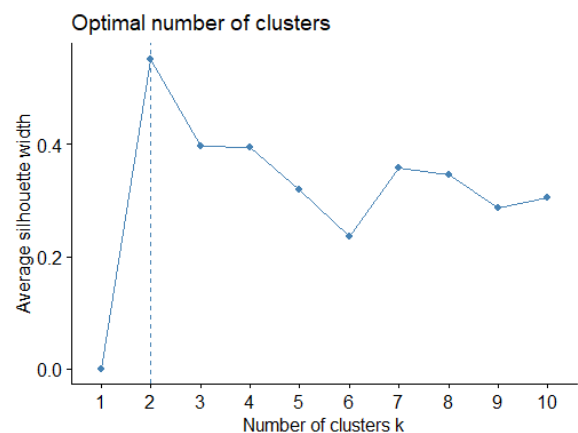
Kota Pekalongan	74,51	12,86	9,20	13158,00
Kota Tegal	74,64	13,08	9,00	13455,00

Data yang diperoleh, selanjutnya diproses menggunakan perhitungan k-means berbasis website untuk membentuk *cluster* berdasarkan indikator IPM di Provinsi Jawa Tengah tahun 2022. Pada statistik deskriptif dibawah ini menunjukkan bahwa semua data yaitu UHH, HLS, RLS, dan PRP pada 35 Kabupaten di Provinsi Jawa Tengah

Tabel 2. Statistik Deskriptif

	N	Minimum	Maximum	Mean	Std. Deviation
Usia Harapan Hidup	35	69.74	77.82	75.1866	1.73224
Harapan Lama Sekolah	35	11.78	15.54	13.0266	.91197
Rata - Rata Lama Sekolah	35	6.35	10.95	8.1411	1.26889
Pengeluaran per Kapita Disesuaikan	35	8994	16351	11533.60	1813.539
Valid N (listwise)	35				

Selanjutnya menentukan cluster, karena analisis K-Means merupakan analisis non hierarki. Maka nilai k ditentukan sendiri oleh peneliti, namun untuk membuktikan nilai k yang digunakan paling optimal dapat menggunakan metode *Silhouette* sebagai berikut.

Gambar 2. Analisis Jumlah Clustering K-Means dengan *Silhouette*

Untuk penentuan nilai K pada metode silhouette, dilihat dari garis tertinggi atau melihatnya dengan melihat garis yang paling optimum adalah 2, namun ada opsional lain yaitu menentukan nilai K dengan menggunakan grafik paling tinggi berikutnya setelah grafik paling tinggi pertama yakni grafik tertinggi kedua, yaitu di angka 3 dan 4, jadi penulis memutuskan nilai K optimum metode silhouette adalah 4.

Setelah menentukan jumlah *cluster* dan diketahui 35 Kabupaten / Kota di Provinsi Jawa Tengah, dengan menggunakan data yang telah distandarisasi dilakukan pembentukan *cluster*. Berdasarkan Gambar

Tabel 2 dengan ketentuan yang telah dijelaskan, dapat didefinisikan sebagai berikut.

Tabel 3. *Cluster Means*

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Usia Harapan Hidup Saat Lahir	1.31531	0.35490	-2.53520	-0.39388
Harapan Lama Sekolah	2.21053	0.15435	-0.34713	-0.74736
Rata-Rata Lama Sekolah	2.17621	0.22870	-1.05695	-0.71581
Pengeluaran per Kapita Disesuaikan	2.00473	0.26699	-0.69841	-0.75907

Berdasarkan Tabel 3 diatas dijelaskan variabel Usia Harapan Hidup Saat Lahir, Harapan Lama Sekolah, Rata-Rata Lama Sekolah, Pengeluaran per Kapita Disesuaikan pada Cluster 1 dan 2 mempunyai rata-rata yang diatas total, sedangkan pada cluster 3 dan 4 mempunyai rata-rata yang dibawah normal karena bernilai negatif. Setelah membentuk 4 *cluster*, langkah selanjutnya adalah melihat apakah variabel-variabel yang telah terbentuk *cluster* mempunyai perbedaan pada setiap *cluster*. Dalam hal ini dapat dilihat dari *F* dan nilai probabilitas (*sig*) masing-masing variabel. Hal ini dilakukan dengan melihat *output* Anova berikut.

Tabel 4 Anova

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	Df		
Z_UHH	7.945	3	0.328	31	24.233	0
Z_HLS	9.321	3	0.195	31	47.872	0
Z_RLS	9.712	3	0.157	31	61.895	0
Z_PRP	8.729	3	0.252	31	34.635	0

Berdasarkan Tabel 4 diatas jika semakin besar angka *F* suatu variabel dan angka *sig* adalah 0.05 maka semakin besar pula perbedaan variabel tersebut pada variabel lainnya. Sebagai contoh, angka *F* terbesar (61,895) ada pada Z RLS, dengan angka *sig* 0,000 yang berarti signifikansinya adalah nyata. Hal ini berarti faktor angka RLS sangat membedakan karakteristik empat *cluster* tersebut. atau dapat dikatakan angka RLS pada keempat *cluster* yang sangat berbeda antara *cluster* 1 dengan *cluster* lainnya. Begitu juga dengan variabel lainnya. Untuk mengetahui lebih detail Kabupaten/Kota masing-masing *cluster* dapat dilihat pada Tabel 5

Tabel 5. Anggota Clustering

Anggota Clustering					
1	Jarak	2	Jarak	3	Jarak
Kota Magelang	1.548	Kabupaten Banyumas	1.225	Kabupaten Tegal	0.751
Kota Surakarta	0.234	Kabupaten Purworejo	0.953	Kabupaten Brebes	0.751
Kota Salatiga	0.793	Kabupaten Boyolali	0.966	Kabupaten Banjarnegara	0.700
Kota Semarang	0.745	Kabupaten Klaten	0.923	Kabupaten Kebumen	1.387
		Kabupaten Sukoharjo	1.700	Kabupaten Wonosobo	0.914
		Kabupaten Karanganyar	1.250	Kabupaten Magelang	0.596
		Kabupaten Sragen	0.814	Kabupaten Wonogiri	1.144
		Kabupaten Pati	0.866	Kabupaten Grobogan	0.369
		Kabupaten Kudus	0.639	Kabupaten Blora	0.217
		Kabupaten Jepara	0.802	Kabupaten Rembang	0.520
		Kabupaten Demak	0.810	Kabupaten Temanggung	0.769
		Kabupaten Semarang	0.409	Kabupaten Batang	0.396
		Kabupaten Kendal	0.953	Kabupaten Pekalongan	0.546
		Kota Pekalongan	1.196	Kabupaten Pemalang	1.072
		Kota Tegal	1.135		

Pada Tabel 5 menunjukkan nomor *cluster* dari Kabupaten/Kota, dan kolom jarak ke pusat *cluster* menunjukkan jarak terdekat antara data dengan pusat *cluster*, dimana semakin kecil berarti semakin mirip dengan *cluster* terkait. Jarak antar *cluster* dapat dilihat dari output *Distances between Final Cluster Centers* yang telah dirangkum dalam Tabel 6

Tabel 6 Jarak Antar Cluster

Cluster	1	2	3	4
1		3.459	6.255	5.260
2	3.459		3.345	1.822
3	6.255	3.345		2.206
4	5.260	1.822	2.206	

*Distance-between* atau *centroid-distance*, adalah jarak antar *centroid* dari satu *cluster* ke *cluster* lainnya. Hasil *Clustering* dikatakan baik jika memiliki nilai *distance between* yang relatif tinggi. Semakin besar nilainya, berarti jarak antara *cluster* satu dengan lainnya semakin renggang. Hal ini berarti antara *cluster* satu dengan lainnya terlihat semakin jelas perbedaannya.

#### 4. Kesimpulan dan Saran

Berdasarkan hasil penelitian yang dilakukan dengan K-Means *Cluster* pada data indikator IPM Kabupaten/Kota di Provinsi Jawa Tengah Tahun 2022, diperoleh kesimpulan pengelompokan Kabupaten/Kota berdasarkan indikator Usia Harapan Hidup (UHH), Harapan Lama Sekolah (HLS), Rata-Rata Lama Sekolah (RLS), dan Pengeluaran Perkapita (PRP) sebagai berikut:

- 1) *Cluster* 1 yaitu Kota Magelang, Kota Surakarta, Kota Salatiga, dan Kota Semarang. Hal ini dikarenakan karakteristik Kabupaten/Kota di *cluster* 1 memiliki kesamaan karakteristik yaitu nilai semua indikator IPM berada di kategori tinggi jika dibandingkan dengan nilai indikator pada *cluster* lain.
- 2) *Cluster* 2 yaitu Kabupaten Banyumas, Kabupaten Purworejo, Kabupaten Boyolali, Kabupaten Klaten, Kabupaten Sukoharjo, Kabupaten Karanganyar, Kabupaten Sragen, Kabupaten Pati, Kabupaten Kudus, Kabupaten Jepara, Kabupaten Demak, Kabupaten Semarang, Kabupaten Kendal, Kota Pekalongan, dan Kota Tegal. Hal ini dikarenakan karakteristik Kabupaten/Kota di *cluster* 2 memiliki kesamaan karakteristik yaitu nilai semua indikator IPM berada di kategori menengah jika dibandingkan dengan nilai indikator pada *cluster* lain.
- 3) *Cluster* 3 yaitu Kabupaten Tegal dan Kabupaten Brebes. Hal ini dikarenakan karakteristik Kabupaten/Kota di *cluster* 3 memiliki kesamaan karakteristik yaitu nilai semua indikator IPM berada kategori rendah jika dibandingkan dengan nilai indikator pada *cluster* lain.
- 4) *Cluster* 4 yaitu Kabupaten Cilacap, Kabupaten Purbalingga, Kabupaten Banjarnegara, Kabupaten Kebumen, Kabupaten Wonosobo, Kabupaten Magelang, Kabupaten Wonogiri, Kabupaten Grobogan, Kabupaten Blora, Kabupaten Rembang, Kabupaten Temanggung, Kabupaten Batang, Kabupaten Pekalongan, dan Kabupaten Pemalang. Hal ini dikarenakan karakteristik Kabupaten/Kota di *cluster* 4 memiliki kesamaan karakteristik yaitu nilai semua indikator IPM berada kategori sangat rendah jika dibandingkan dengan nilai indikator pada *cluster* lain.

Dengan kesimpulan yang disampaikan diatas, peneliti bermaksud memberikan saran – saran sebagai berikut:

- 1) Pada penelitian ini peneliti hanya mengkaji Indeks Pembangunan Manusia menggunakan klasifikasi UHH ( Usia Harapan Hidup), HLS (Harapan Lama Sekolah), RLS (Rata – rata Lama Sekolah), PKP (Pendapatan per Kapita. Oleh karena itu angka Harapan Hidup, dan lain sebagainya
- 2) Hasil penelitian ini dapat digunakan sebagai salah satu bahan untuk peneliti lanjutan, mungkin bisa menggunakan metode clustering lainnya, selain Metode K – Means yang digunakan dalam penelitian ini, mengingat metode untuk clustering sangatlah beragam. Sehingga pembaca dapat mengkaji data Indeks Pembangunan Manusia (IPM) Menggunakan metode hierarki yang lainnya.

#### 5. Daftar Pustaka

- [1] Anggraini, L. and Arum, P.R., 2022. Analisis Cluster Menggunakan Algoritma K-Means Pada Provinsi Sumatera Barat Berdasarkan Indeks Pembangunan Manusia Tahun 2021. In *Prosiding Seminar Nasional UNIMUS* (Vol. 5).
- [2] Anwar, K., Goejantoro, R. and Prangga, S., 2022. Pengelompokan Kabupaten/Kota Di Pulau Kalimantan Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2020 Menggunakan Optimasi K-Means Cluster Dengan Principle Component Analysis (PCA). *EKSPONENSLAL*, 13(2), pp.131-140. DOI: <https://doi.org/10.30872/eksponensial.v13i2.1053>.
- [3] Sibarani, H., Saputra, W., Gunawan, I. and Nasution, Z.M., 2022. Penerapan Metode K-Means Untuk Pengelompokan Kabupaten/Kota Di Provinsi Sumatera Utara Berdasarkan Indikator Indeks Pembangunan Manusia. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(1), pp.154-161. DOI: <https://doi.org/10.36040/jati.v6i1.4590>

- [4] Kurnia, A., Rahardiantoro, S. and Mattjik, A.A., 2022. Penggerombolan Kabupaten/Kota di Indonesia Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan Metode K-Means dan Fuzzy C-Means. *Xplore: Journal of Statistics*, 11(1), pp.36-47. DOI: <https://doi.org/10.29244/xplore.v11i1.855>
- [5] Hermawan, H. and Hasugian, H., 2022, September. Penerapan Data Mining Untuk Clustering Indeks Pembangunan Manusia Berdasarkan Provinsi Di Indonesia. In *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)* (Vol. 1, No. 1, pp. 525-532).
- [6] Hartono, Bambang. 2013. *Sistem Informasi Manajemen Berbasis Komputer*. Jakarta: Rineka Cipta.
- [7] Kusrini. 2016. *Sistem Pakar*. Yogyakarta: Andi Offisit.
- [8] Kusuma, V. M. 2017. *Implementasi Metode Fuzzy Subtractive Clustering Untuk Pengelompokan Data Potensi Kebakaran Hutan/Lahan*. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer.
- [9] Muslihudin, Muhamad Oktafianto. 2016. *Analisis dan Perancangan Sistem Informasi Menggunakan Model Terstruktur dan UML*. Yogyakarta: Andi.
- [10] Murdick, Robert dkk. 2016. *Sistem Informasi Untuk Manajemen Modern*. Bandung : Informatika.
- [11] Nahampun, Maruli Tua. 2014. *Sistem Pakar Diagnosa Penyakit Tanaman Kelapa Sawit Dengan Metode Dempster Shafer*. Jurnal Pelita Informatika Budi.
- [12] Prasetyo, Eko. 2016. *Data Mining Mengolah Data Menjadi Informasi. Menggunakan Matlab*. Yogyakarta: Penerbit Andi.
- [13] Rikhiana, Esthi Dyah dan Abdul Fadlil. 2013. *Implementasi Sistem Pakar Untuk Mendiagnosa Penyakit Dalam Pada Manusia Menggunakan Metode Dempster Shafer*. Jurnal Serjana Teknik Informatika.
- [14] Sukamto, R. A., & Shalahuddin, M. (2016). *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek*. Bandung: Informatika.
- [15] Suyatno. 2017. *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika Bandung.