

Analisis Perbandingan Pengelompokan Indeks Pembangunan Manusia Indonesia Tahun 2019 dengan Metode *Partitioning* dan *Hierarchical Clustering*

Arina Mana Sikana¹, Arie Wahyu Wijayanto²

Program Studi Komputasi Statistik Politeknik Statistika STIS
Jalan Otto Iskandardinata 64C Jakarta 13330
¹221810195@stis.ac.id
²ariewahyu@stis.ac.id

Abstract

Human Development Index (HDI) is an important indicator in measuring the level of success of the development of the quality of human life. Human Development Index clustering aims to divide the regions into groups based on Human Development Index for the region in 2019. Human Development Index clustering compares Partitioning Clustering and Hierarchical Clustering method to divide Human Development Index Indonesia in 2019. Partitioning Clustering method uses K-Means Clustering algorithm and Hierarchical Clustering method uses Agglomerative Ward Clustering algorithm. The results obtained are the best method for grouping provinces in Indonesia based on Human Development Index in 2019 is K-Means Clustering method with the optimum number of clusters is 6. This method gives Silhouette Score 0,6291, Calinski-Harabasz Index 241,8875, dan Davies-Bouldin Index 0,3038. While the best method for grouping regencies in Indonesia based on Human Development Index in 2019 is K-Means Clustering method with the optimum number of clusters is 6. This method gives Silhouette Score 0,5511, Calinski-Harabasz Index 1525,4007, dan Davies-Bouldin Index 0,5234.

Keywords: *Human Development Index, K-Means Clustering, Agglomerative Ward Clustering*

Abstrak

Indeks Pembangunan Manusia (IPM) merupakan indikator penting dalam pengukuran tingkat keberhasilan pembangunan kualitas hidup manusia. Pengelompokan Indeks Pembangunan Manusia (IPM) bertujuan untuk membagi wilayah-wilayah ke dalam kelompok berdasarkan Indeks Pembangunan Manusia wilayah tersebut tahun 2019. Pengelompokan Indeks Pembangunan Manusia Indonesia tahun 2019 membandingkan metode *Partitioning Clustering* dan *Hierarchical Clustering*. Algoritma *Partitioning Clustering* yang digunakan adalah algoritma *K-Means Clustering*, sedangkan algoritma *Hierarchical Clustering* adalah algoritma *Agglomerative Ward Clustering*. Hasil yang diperoleh adalah metode terbaik untuk pengelompokan provinsi di Indonesia berdasarkan Indeks Pembangunan Manusia untuk tahun 2019 adalah metode *K-Means Clustering* dengan jumlah kluster optimum adalah 6. Metode ini memberikan *Silhouette Score* sebesar 0,6291, *Calinski-Harabasz Index* sebesar 241,8875, dan *Davies-Bouldin Index* sebesar 0,3038. Sedangkan metode terbaik untuk pengelompokan kabupaten/kota di Indonesia berdasarkan Indeks Pembangunan Manusia untuk tahun 2019 adalah metode *K-Means Clustering* dengan jumlah kluster optimum adalah 6. Metode ini memberikan *Silhouette Score* sebesar 0,5511, *Calinski-Harabasz Index* sebesar 1525,4007, dan *Davies-Bouldin Index* sebesar 0,5234.

Kata Kunci : *Indeks Pembangunan Manusia, K-Means Clustering, Agglomerative Ward Clustering*

1. PENDAHULUAN

Indeks Pembangunan Manusia (IPM) merupakan indikator penting dalam pengukuran tingkat keberhasilan pembangunan kualitas hidup manusia. IPM yang diperkenalkan oleh *United Nations Development Programme* (UNDP) pada tahun 1990 merupakan data strategis sebagai ukuran kinerja pemerintah. IPM menjelaskan tingkat akses penduduk terhadap hasil pembangunan dalam memperoleh pendapatan, Kesehatan, Pendidikan, dan sebagainya [1].

Pengelompokan Indeks Pembangunan Manusia (IPM) bertujuan untuk membagi wilayah-wilayah ke dalam beberapa kelompok yang memiliki tingkat keseragaman karakteristik IPM yang tinggi dalam setiap kelompok dan keberagaman karakteristik IPM yang tinggi antar kelompok [2]. Pengelompokan IPM akan membagi wilayah-wilayah ke dalam kelompok berdasarkan Indeks Pembangunan Manusia wilayah tersebut pada tahun 2019.

Hasil pengelompokan dapat digunakan sebagai bahan evaluasi kinerja pemerintah dalam pembangunan kualitas hidup manusia. Hasil pengelompokan IPM juga dapat digunakan sebagai bahan perencanaan kebijakan pemerintah selanjutnya [3][4]. Wilayah-wilayah dalam kelompok yang sama memiliki kemiripan yang tinggi pada Indeks Pembangunan Manusia, sehingga kebijakan yang diterapkan pada wilayah dalam suatu kelompok dapat dijadikan acuan kebijakan wilayah lain dalam kelompok yang sama.

Pada penelitian ini, pengelompokan Indeks Pembangunan Manusia Indonesia tahun 2019 menggunakan metode *Partitioning* dan *Hierarchical Clustering*. Algoritma *Partitioning Clustering* yang digunakan adalah algoritma *K-Means Clustering*, sedangkan algoritma *Hierarchical Clustering* adalah algoritma *Agglomerative Ward Clustering*. Penelitian ini bertujuan mencari metode terbaik untuk mengelompokkan Indeks Pembangunan Manusia Indonesia tahun 2019.

2. METODE PENELITIAN

2.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari situs Badan Pusat Statistik (bps.go.id). Variabel yang digunakan adalah Indeks Pembangunan Manusia (IPM), Umur Harapan Hidup Saat Lahir (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan Pengeluaran per Kapita Disesuaikan dari 34 provinsi dan 514 kabupaten/kota di Indonesia tahun 2019.

2.2 Metode *Partitioning Clustering*

Algoritma *clustering* yang paling terkenal dan sering digunakan adalah algoritma *partitioning*. Algoritma ini lebih sederhana, mudah diterapkan, dan memiliki kompleksitas waktu yang lebih kecil dibandingkan teknik lain [5][6]. Algoritma *partitioning* yang akan dianalisis adalah algoritma *K-Means*.

Algoritma *K-Means* membagi satu set n sampel X ke dalam k kluster yang saling lepas. Algoritma *K-Means* mengelompokkan data dengan memisahkannya ke dalam beberapa kelompok varians yang sama dan meminimalkan kriteria inersia atau jumlah kuadrat dalam kluster. Kriteria inersia dinyatakan dalam persamaan (1) berikut,

$$\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|^2) \quad (1)$$

dengan μ_j adalah rata-rata yang sering disebut sebagai *cluster centroid* [7].

Inersia yang merupakan jumlah kuadrat dalam kluster dapat dijadikan sebagai ukuran seberapa koheren anggota-anggota dalam kluster yang sama. Pada penelitian ini, pembentukan kluster dengan metode *K-Means Cluster* menggunakan fungsi *KMeans* dari *Scikit Learn Clustering* di python.

2.3 Metode *Hierarchical Clustering*

Hierarchical Cluster melakukan pengelompokan dengan membangun kluster bersarang dengan menggabungkan atau memisahkannya secara berurutan. *Hierarchical Cluster* direpresentasikan sebagai pohon atau dendrogram yang daunnya menjadi kelompok hanya

dengan satu anggota [7]. Pada penelitian ini, pembentukan kluster dengan metode *Hierarchical Clustering* menggunakan algoritma *Agglomerative Ward* yaitu menggunakan fungsi *AgglomerativeClustering* dengan kriteria keterkaitan *ward* dari *Scikit Learn Clustering* di python.

Agglomerative Clustering melakukan pengelompokan hirarki dengan pendekatan *bottom-up*, yaitu setiap observasi dimulai di klusternya sendiri dan kluster secara berturut-turut digabungkan. Kriteria keterkaitan *ward* akan meminimalkan perbedaan jumlah kuadrat dalam setiap kluster [7][8].

2.4 Silhouette Coefficient

Silhouette Score merupakan salah satu cara untuk evaluasi performa kluster. *Silhouette Score* dinyatakan dalam persamaan (2) berikut,

$$s = \frac{b - a}{\max(a, b)} \quad (2)$$

dengan a adalah jarak rata-rata antara observasi dengan semua titik lain dalam kluster yang sama dan b adalah jarak rata-rata antara observasi dengan semua titik lain dalam kluster terdekat [7].

Silhouette Score bernilai antara -1 untuk pengelompokan yang salah hingga 1 untuk hasil pengelompokan yang sangat padat. Nilai di sekitar 0 mengindikasikan kluster yang tumpang tindih. *Silhouette Score* yang lebih tinggi mengindikasikan model dengan kluster yang lebih jelas atau model memiliki performa yang lebih baik [7][6]. ada penelitian ini, penghitungan *Silhouette Score* menggunakan fungsi *silhouette_score* dari *Scikit Learn Metrics* di python.

2.5 Calinski-Harabasz Index

Calinski-Harabasz Index yang juga dikenal sebagai *Variance Ratio Criterion*, dapat digunakan untuk evaluasi model *clustering*. *Calinski-Harabasz Index* adalah rasio jumlah dispersi antar kluster dan dispersi dalam kluster untuk semua kluster (dispersi didefinisikan sebagai jumlah kuadrat jarak) [7][9][10]. Untuk sekumpulan data X berukuran n yang telah dikelompokkan menjadi k kluster, *Calinski-Harabasz Index* dinyatakan dalam persamaan (3) berikut,

$$s = \frac{tr(B_k)}{tr(W_k)} \times \frac{n - k}{k - 1} \quad (3)$$

dengan $tr(B_k)$ adalah *trace* dari matriks disperse antar kluster dan $tr(W_k)$ adalah *trace* matriks disperse dalam kluster yang dinyatakan dalam persamaan (4) dan (5) berikut,

$$B_k = \sum_{q=1}^k n_q (c_q - c_x)(c_q - c_x)^T \quad (4)$$

$$W_k = \sum_{q=1}^k \sum_{x \in C_q} (x - c_q)(x - c_q)^T \quad (5)$$

dengan C_q adalah kumpulan titik dalam kluster q , c_q adalah pusat kluster q , c_x adalah pusat data X , dan n_q adalah jumlah titik dalam kluster q . *Calinski-Harabasz Index* yang lebih tinggi berkaitan dengan model dengan kluster yang lebih jelas [7]. Pada penelitian ini, penghitungan *Calinski-Harabasz Index* menggunakan fungsi *calinski_harbasz_score* dari *Scikit Learn Metrics* di python.

2.6 Davies-Bouldin Index

Davies-Bouldin Index menyatakan rata-rata 'kesamaan' antar kluster yang merupakan ukuran yang membandingkan jarak antar kluster dengan ukuran kluster itu sendiri [7][11][10]. 'Kesamaan' yang didefinisikan sebagai ukuran R_{ij} dinyatakan dalam persamaan (6) berikut,

$$R_{ij} = \frac{s_i + s_j}{d_{ij}} \quad (6)$$

dengan s_i adalah rata-rata jarak tiap titik dalam kluster i dan pusat kluster I , juga dikenal sebagai diameter kluster dan d_{ij} adalah jarak antara pusat kluster i dan pusat kluster j . *Davies-Bouldin Index* dinyatakan dalam persamaan (7) berikut,

$$s = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} R_{ij}$$

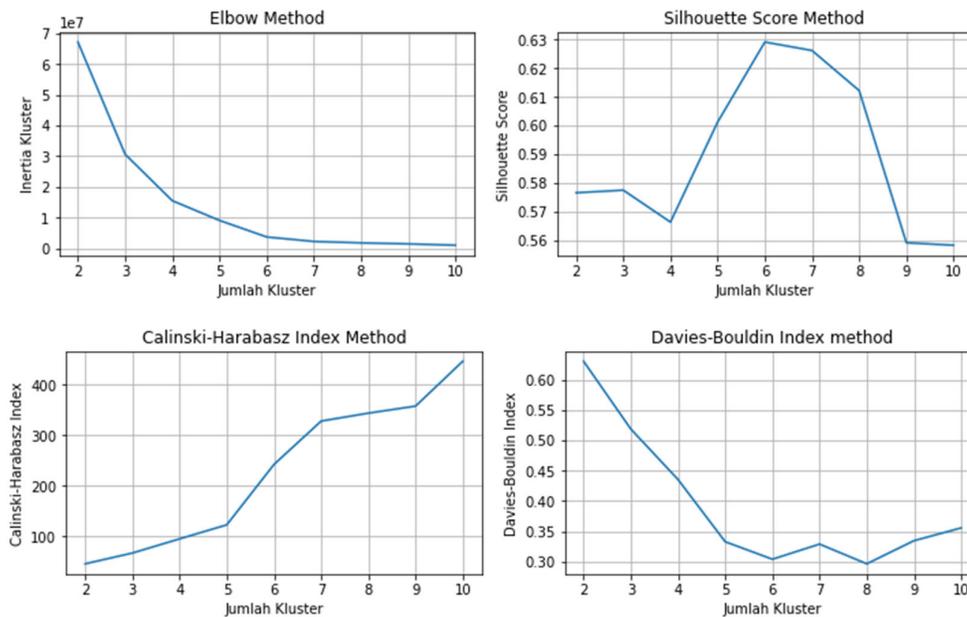
dengan R_{ij} merupakan ukuran ‘kesamaan’ dari persamaan (6). Nilai *Davies-Bouldin Index* terkecil adalah 0. *Davies-Bouldin Index* yang mendekati 0 menunjukkan partisi yang lebih baik [7]. Pada penelitian ini, penghitungan *Davies-Bouldin Index* menggunakan fungsi *davies_bouldin_score* dari *Scikit Learn Metrics* di python.

3. HASIL DAN PEMBAHASAN

3.1 Pengelompokan IPM Provinsi

3.1.1 K-Means Clustering

Penentuan jumlah kluster optimum untuk pengelompokan IPM Provinsi dengan metode *K-Means Clustering* menggunakan *Elbow Method*, *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index*. Berikut adalah grafik *Elbow Method*, *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index* untuk setiap jumlah kluster untuk IPM provinsi yang akan dibentuk menggunakan metode *K-Means Clustering*.

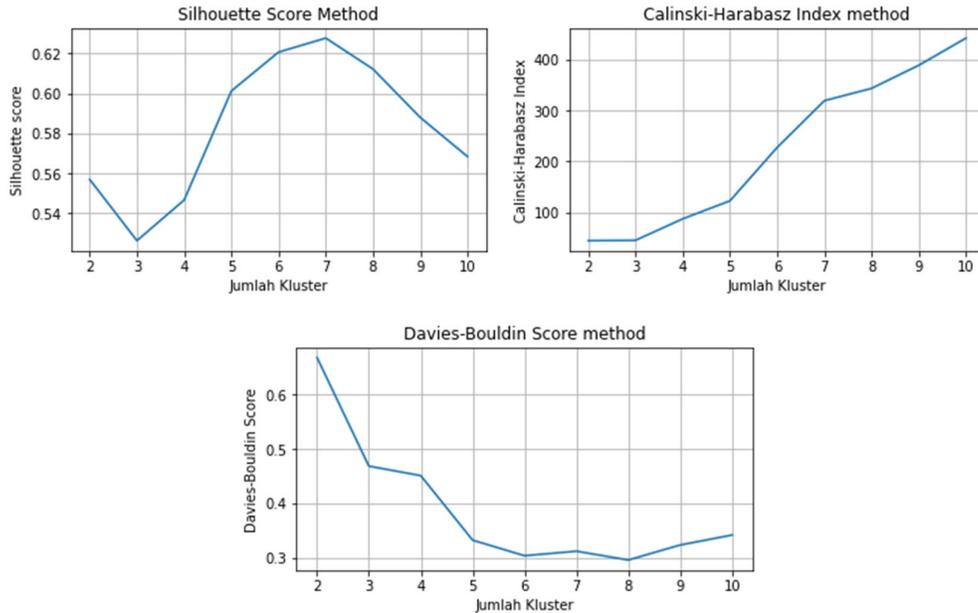


Gambar 1. *Elbow Method*, *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index K-Means Clustering* IPM Provinsi

Dari *Elbow Method* pada Gambar 1 diatas, garis membentuk bengkokan (*elbow*) pada jumlah kluster 3, 4, dan 6. Sedangkan untuk *Silhouette Score*, *Silhouette Score* tertinggi terdapat pada jumlah kluster 6. Pada *Calinski-Harabasz Index*, *Calinski-Harabasz Index* terus meningkat seiring bertambahnya jumlah kluster. Pada *Davies-Bouldin Index*, *Davies-Bouldin Index* terendah terdapat pada jumlah kluster 6 dan 8. Dari Gambar 1 diatas, dapat disimpulkan bahwa pengelompokan IPM provinsi menggunakan metode *K-Means Clustering* akan optimum jika jumlah kluster adalah 6.

3.1.2 Agglomerative Ward Clustering

Penentuan jumlah kluster optimum untuk pengelompokan IPM provinsi dengan metode *Agglomerative Ward Clustering* menggunakan *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index*. Berikut adalah grafik nilai *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index* untuk setiap jumlah kluster untuk IPM provinsi yang akan dibentuk menggunakan metode *Agglomerative Ward Clustering*.



Gambar 2. *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index* *Agglomerative Ward Clustering* IPM Provinsi

Dari *Silhouette Score* pada Gambar 2, jumlah kluster dengan *Silhouette Score* tertinggi adalah 7. Pada *Calinski-Harabasz Index*, *Calinski-Harabasz Index* terus meningkat seiring bertambahnya jumlah kluster. Pada *Davies-Bouldin Index*, *Davies-Bouldin Index* terendah terdapat pada jumlah kluster 6 dan 8. Namun *Davies-Bouldin Index* juga memiliki nilai yang kecil pada jumlah kluster 7, sehingga dapat disimpulkan bahwa pengelompokan IPM provinsi menggunakan metode *Agglomerative Ward Clustering* akan optimum jika jumlah kluster adalah 7.

3.1.3 Metode Terbaik

Pemilihan metode terbaik untuk pengelompokan IPM provinsi dilakukan dengan membandingkan *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index* antara *K-Means Clustering* dengan jumlah kluster optimum adalah 6 dan *Agglomerative Ward Clustering* dengan jumlah kluster optimum adalah 7. *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index* metode *K-Means Clustering* dan *Agglomerative Ward Clustering* disajikan pada Tabel 1 berikut.

Tabel 1. Evaluasi Model *K-Means Clustering* dan *Agglomerative Ward Clustering* IPM Provinsi

Evaluasi Model	<i>K-Means Clustering</i>	<i>Agglomerative Ward Clustering</i>
<i>Silhouette Score</i>	0,6291	0,6276
<i>Calinski-Harabasz Index</i>	241,887	319,277

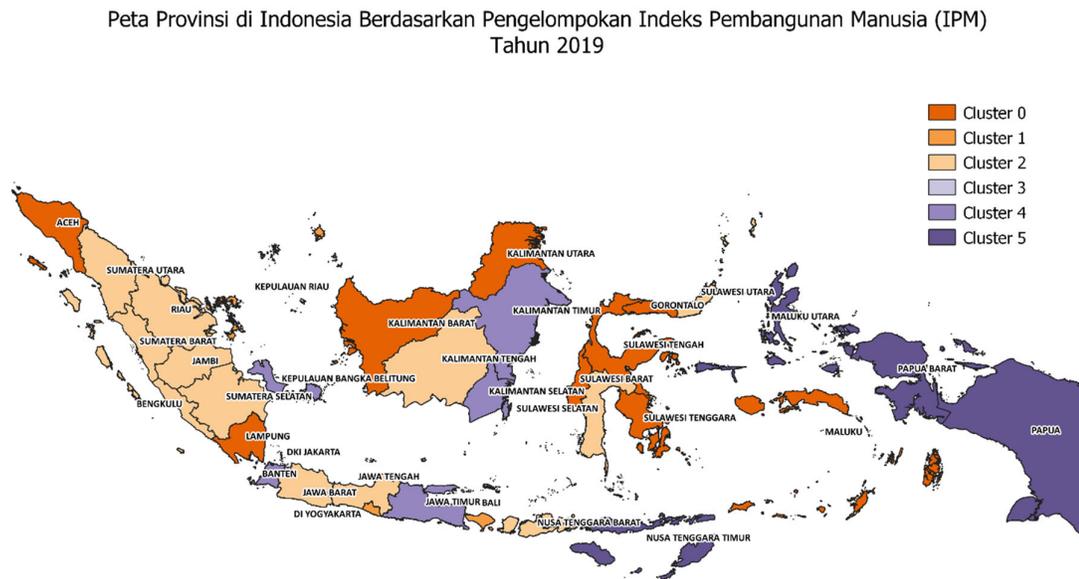
<i>Davies-Bouldin Index</i>	0,3038	0,3124
-----------------------------	--------	--------

Dari Tabel 1 diatas, *Silhouette Score* metode *K-Means Clustering* lebih tinggi dibandingkan metode *Agglomerative Ward Clustering*, *Calinski-Harabasz Index* metode *K-Means Clustering* lebih kecil dibandingkan metode *Agglomerative Ward Clustering*, dan *Davies-Bouldin Index* metode *K-Means Clustering* lebih kecil dibandingkan metode *Agglomerative Ward Clustering*. Hal ini menunjukkan bahwa berdasarkan *Silhouette Score* dan *Davies-Bouldin Index*, metode *K-Means Clustering* dengan jumlah kluster optimum adalah 6 lebih baik dan akan digunakan untuk mengelompokkan IPM provinsi di Indonesia tahun 2019.

Dari kelompok-kelompok yang sudah terbentuk, selanjutnya akan dilihat karakteristik tiap kelompok untuk setiap variabel penelitian. Karakteristik tiap kelompok ditentukan dengan membandingkan antara rata-rata masing-masing kelompok terhadap rata-rata keseluruhan untuk setiap variabel penelitian.

Kelompok yang memiliki rata-rata kelompok lebih dari rata-rata keseluruhan untuk setiap variabel diberi kode 1 dan kelompok yang memiliki rata-rata kelompok kurang dari rata-rata keseluruhan untuk setiap variabel diberi kode 0. Hasil pengelompokan provinsi berdasarkan Indeks Pembangunan Manusia tahun 2019 disajikan dalam Gambar 3 berikut. Karakteristik tiap kelompok untuk setiap variabel penelitian disajikan dalam Tabel 2 berikut.

Dari Tabel 2 dibawah, kelompok 3 memiliki rata-rata kelompok untuk HLS kurang dari rata-rata seluruh provinsi, sedangkan rata-rata kelompok untuk variabel lain lebih dari rata-rata seluruh provinsi. Kelompok 3 hanya memiliki 1 anggota, yaitu provinsi DKI Jakarta



Gambar 3. Peta Provinsi di Indonesia Berdasarkan Pengelompokan Indeks Pembangunan Manusia Tahun 2019

Tabel 2. Karakteristik Kelompok Provinsi Berdasarkan Rata-Rata

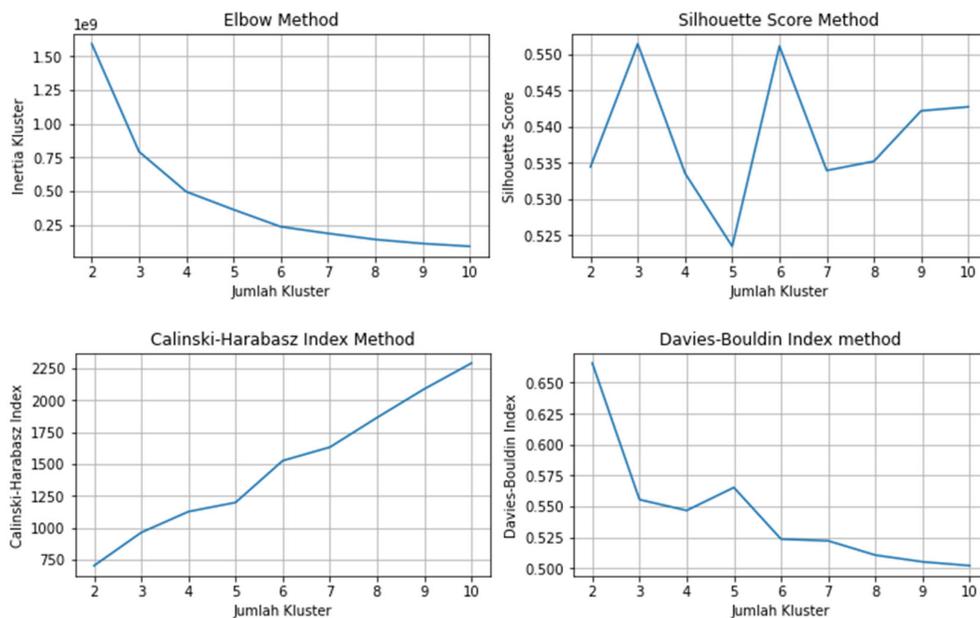
Kelompok	IPM	UHH	RLS	HLS	Pengeluaran	Jumlah Anggota
0	0	0	0	1	0	9
1	1	1	1	1	1	3

2	1	1	0	0	1	12
3	1	1	1	0	1	1
4	1	1	0	0	1	5
5	0	0	0	0	0	4

3.2 Pengelompokan IPM Kabupaten/Kota

3.2.1 K-Means Clustering

Penentuan jumlah kluster optimum untuk pengelompokan IPM kabupaten/kota dengan metode *K-Means Clustering* menggunakan *Elbow Method*, *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index*. Berikut adalah grafik *Elbow Method*, *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index* untuk setiap jumlah kluster untuk IPM kabupaten/kota yang akan dibentuk menggunakan metode *K-Means Clustering*.



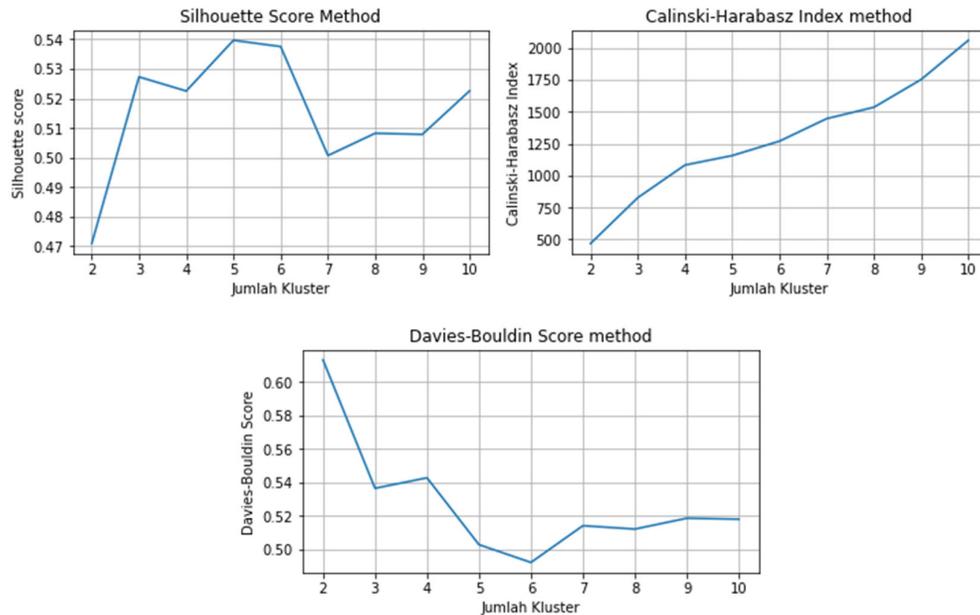
Gambar 4. *Elbow Method*, *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index* *K-Means Clustering* IPM Kabupaten/Kota

Dari *Elbow Method* pada Gambar 4 diatas, garis membentuk bengkokan (*elbow*) pada jumlah kluster 3, 4, dan 6. Sedangkan untuk *Silhouette Score*, *Silhouette Score* tertinggi terdapat pada jumlah kluster 3 dan 6. Pada *Calinski-Harabasz Index*, *Calinski-Harabasz Index* terus meningkat seiring bertambahnya jumlah kluster. Pada *Davies-Bouldin Index*, *Davies-Bouldin Index* cenderung mengalami penurunan seiring bertambahnya jumlah kluster, kecuali pada jumlah kluster 4. Dari Gambar 4 diatas, dapat disimpulkan bahwa pengelompokan IPM kabupaten/kota menggunakan metode *K-Means Clustering* akan optimum jika jumlah kluster adalah 6.

3.2.2 Agglomerative Ward Clustering

Penentuan jumlah kluster optimum untuk pengelompokan IPM kabupaten/kota dengan metode *Agglomerative Ward Clustering* menggunakan *Silhouette Score*, *Calinski-Harabasz Index*, dan *Davies-Bouldin Index*. Berikut adalah grafik nilai *Silhouette Score*, *Calinski-Harabasz Index*,

dan *Davies-Bouldin Index* untuk setiap jumlah kluster untuk IPM kabupaten/kota yang akan dibentuk menggunakan metode *Agglomerative Ward Clustering*.



Gambar 5. *Silhouette Score, Calinski-Harabasz Index, dan Davies-Bouldin Index Agglomerative Ward Clustering IPM Kabupaten/Kota*

Dari *Silhouette Score* pada Gambar 5, jumlah kluster dengan *Silhouette Score* tertinggi adalah 5. Pada *Calinski-Harabasz Index*, *Calinski-Harabasz Index* terus meningkat seiring bertambahnya jumlah kluster. Pada *Davies-Bouldin Index*, *Davies-Bouldin Index* terendah terdapat pada jumlah kluster 6. Namun *Davies-Bouldin Index* juga memiliki nilai yang kecil pada jumlah kluster 5, sehingga dapat disimpulkan bahwa pengelompokan IPM provinsi menggunakan metode *Agglomerative Ward Clustering* akan optimum jika jumlah kluster adalah 5.

3.2.3 Metode Terbaik

Pemilihan metode terbaik untuk pengelompokan IPM kabupaten/kota dilakukan dengan membandingkan *Silhouette Score, Calinski-Harabasz Index, dan Davies-Bouldin Index* antara *K-Means Clustering* dengan jumlah kluster optimum adalah 6 dan *Agglomerative Ward Clustering* dengan jumlah kluster optimum adalah 5. *Silhouette Score, Calinski-Harabasz Index, dan Davies-Bouldin Index* metode *K-Means Clustering* dan *Agglomerative Ward Clustering* disajikan pada Tabel 3 berikut.

Tabel 3. Evaluasi Model *K-Means Clustering* dan *Agglomerative Ward Clustering* IPM Kabupaten/Kota

Evaluasi Model	<i>K-Means Clustering</i>	<i>Agglomerative Ward Clustering</i>
<i>Silhouette Score</i>	0,5511	0,5396
<i>Calinski-Harabasz Index</i>	1525,401	1156,555
<i>Davies-Bouldin Index</i>	0,5234	0,5027

Dari Tabel 3 diatas, *Silhouette Score* metode *K-Means Clustering* lebih tinggi dibandingkan metode *Agglomerative Ward Clustering*, *Calinski-Harabasz Index* metode *K-Means Clustering* lebih tinggi dibandingkan metode *Agglomerative Ward Clustering*, dan *Davies-Bouldin Index* metode *K-Means Clustering* lebih tinggi dibandingkan metode *Agglomerative Ward Clustering*. Hal ini menunjukkan bahwa berdasarkan *Silhouette Score* dan *Calinski-Harabasz Index* metode *K-Means Clustering* dengan jumlah kluster optimum adalah 6 lebih baik dan akan digunakan untuk mengelompokkan IPM kabupaten/kota di Indonesia tahun 2019.

Dari kelompok-kelompok yang sudah terbentuk, selanjutnya akan dilihat karakteristik tiap kelompok untuk setiap variabel penelitian. Karakteristik tiap kelompok ditentukan dengan membandingkan antara rata-rata masing-masing kelompok terhadap kuantil data keseluruhan untuk setiap variabel penelitian.

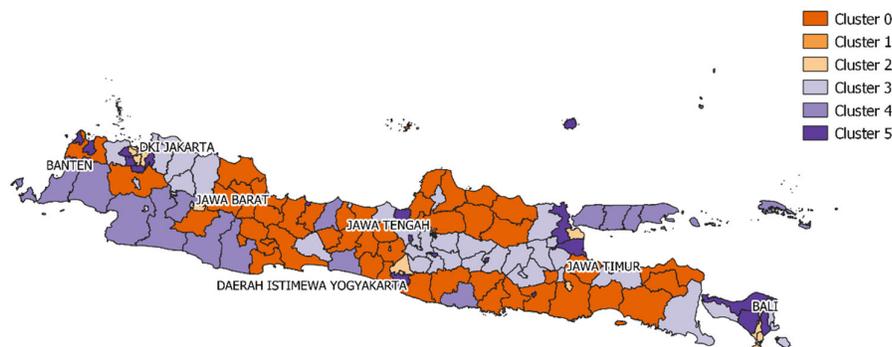
Kelompok yang memiliki rata-rata kelompok kurang dari kuantil bawah untuk setiap variabel diberi kode 1, kelompok yang memiliki rata-rata kelompok antara kuantil bawah dan median untuk setiap variabel diberi kode 2, kelompok yang memiliki rata-rata kelompok antara median dan kuantil atas untuk setiap variabel diberi kode 3, dan kelompok yang memiliki rata-rata kelompok lebih dari kuantil atas untuk setiap variabel diberi kode 4.

Karakteristik tiap kelompok untuk setiap variabel penelitian dan jumlah anggota tiap kelompok disajikan dalam Tabel 4 berikut. Hasil pengelompokan kabupaten/kota berdasarkan Indeks Pembangunan Manusia tahun 2019 disajikan dalam Gambar 8, Gambar 9, Gambar 10, Gambar 11, Gambar 12, Gambar 13, serta Tabel 5 Hasil Pengelompokan Kabupaten/Kota di Indonesia Berdasarkan Indeks Pembangunan Manusia Tahun 2019.

Tabel 4. Karakteristik Kelompok Kabupaten/Kota Berdasarkan Kuantil

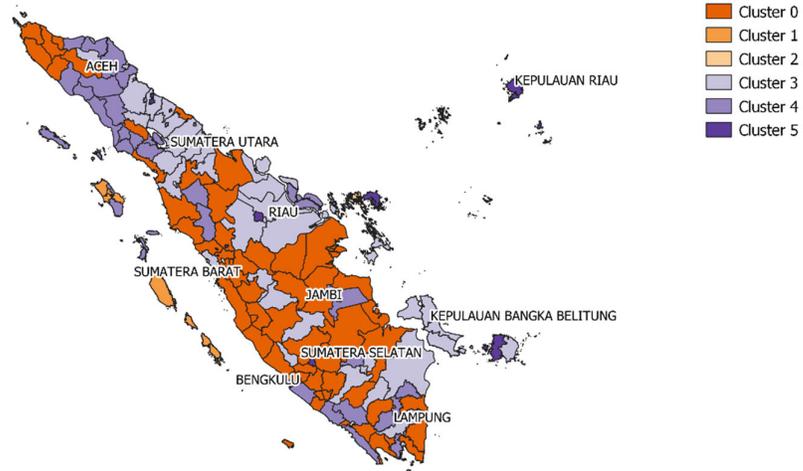
Kelompok	IPM	UHH	RLS	HLS	Pengeluaran	Jumlah Anggota
0	2	2	2	2	2	153
1	1	1	1	1	1	42
2	4	4	4	4	4	16
3	3	3	3	3	4	134
4	1	2	2	2	1	120
5	4	4	4	4	4	49

Peta Kabupaten/Kota di Pulau Jawa dan Bali Berdasarkan Pengelompokan Indeks Pembangunan Manusia (IPM) Tahun 2019



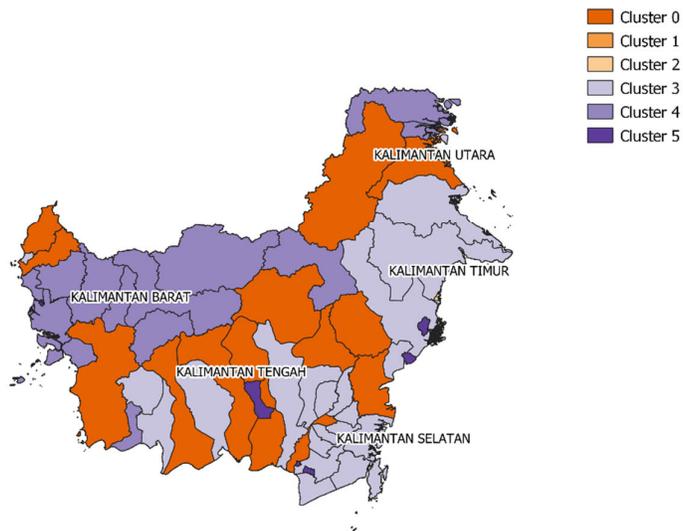
Gambar 8. Peta Kabupaten/Kota di Pulau Jawa dan Bali Berdasarkan Pengelompokan Indeks Pembangunan Manusia Tahun 2019

Peta Kabupaten/Kota di Pulau Sumatera Berdasarkan Pengelompokan Indeks Pembangunan Manusia (IPM) Tahun 2019



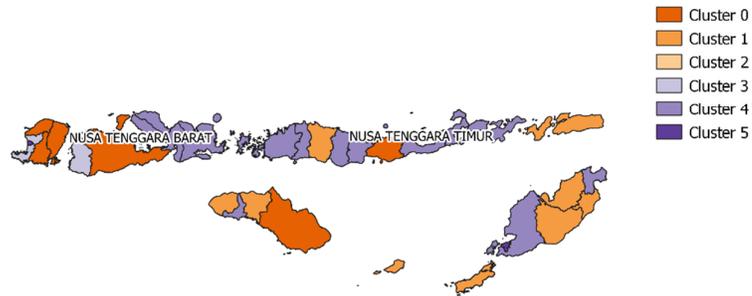
Gambar 9. Peta Kabupaten/Kota di Pulau Sumatera Berdasarkan Pengelompokan Indeks Pembangunan Manusia Tahun 2019

Peta Kabupaten/Kota di Pulau Kalimantan Berdasarkan Pengelompokan Indeks Pembangunan Manusia (IPM) Tahun 2019



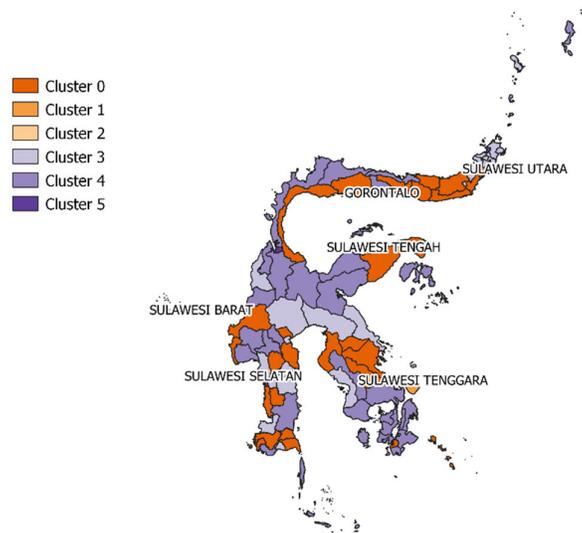
Gambar 10. Peta Kabupaten/Kota di Pulau Kalimantan Berdasarkan Pengelompokan Indeks Pembangunan Manusia Tahun 2019

Peta Kabupaten/Kota di Wilayah Nusa Tenggara Berdasarkan Pengelompokan Indeks Pembangunan Manusia (IPM) Tahun 2019



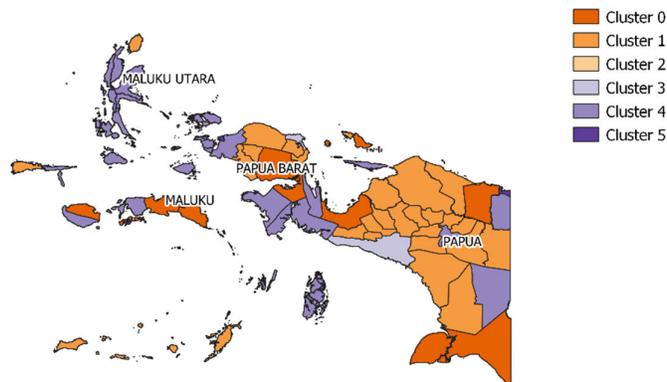
Gambar 11. Peta Kabupaten/Kota di Wilayah Nusa Tenggara Berdasarkan Pengelompokan Indeks Pembangunan Manusia Tahun 2019

Peta Kabupaten/Kota di Pulau Jawa dan Bali Berdasarkan Pengelompokan Indeks Pembangunan Manusia (IPM) Tahun 2019



Gambar 12. Peta Kabupaten/Kota di Pulau Sulawesi Berdasarkan Pengelompokan Indeks Pembangunan Manusia Tahun 2019

Peta Kabupaten/Kota di Maluku dan Papua Berdasarkan Pengelompokan Indeks Pembangunan Manusia (IPM) Tahun 2019



Gambar 13. Peta Kabupaten/Kota di Maluku dan Papua Berdasarkan Pengelompokan Indeks Pembangunan Manusia Tahun 2019

4. KESIMPULAN

Berdasarkan hasil dan pembahasan diatas, dapat disimpulkan bahwa :

1. Metode terbaik untuk pengelompokan provinsi di Indonesia berdasarkan Indeks Pembangunan Manusia untuk tahun 2019 adalah metode *K-Means Cluster* dengan jumlah kluster optimum adalah 6. Metode ini memberikan *Silhouette Score* sebesar 0,6291, *Calinski-Harabasz Index* sebesar 241,8875, dan *Davies-Bouldin Index* sebesar 0,3038.
2. Metode terbaik untuk pengelompokan kabupaten/kota di Indonesia berdasarkan Indeks Pembangunan Manusia untuk tahun 2019 adalah metode *K-Means Cluster* dengan jumlah kluster optimum adalah 6. Metode ini memberikan *Silhouette Score* sebesar 0,5511, *Calinski-Harabasz Index* sebesar 1525,4007, dan *Davies-Bouldin Index* sebesar 0,5234.

DAFTAR PUSTAKA

- [1] BPS, "Indeks Pembangunan Manusia," 2020. <https://bps.go.id/subject/26/indeks-pembangunan-manusia.html> (accessed Nov. 25, 2020).
- [2] A. S. Rizal and R. F. Hakim, "Metode K-Means Cluster Dan Fuzzy C-Means Cluster (Studi Kasus: Indeks Pembangunan Manusia Di Kawasan Indonesia Timur Tahun 2012)," *Pros. Semin. Nas. Mat. dan Pendidik. Mat. UMS 2015*, pp. 643–657, 2015, [Online]. Available: <https://publikasiilmiah.ums.ac.id/xmlui/handle/11617/5803>.
- [3] R. Dewi Kusumah, B. Warsito, and M. Abdul Mukid, "Perbandingan metode k – means dan self organizing map (Studi kasus: pengelompokan kabupaten/kota di jawa tengah berdasarkan indikator indeks pembangunan manusia 2015)," *J. Gaussian, Vol 6 No 3 tahun 2017*, vol. 6, pp. 429–437, 2017.
- [4] Samuel Samuel Ahmad Budi, "Klasterisasi Indeks Pembangunan Manusia (Ipm) Per Kabupaten Di Indonesia Dengan Menggunakan Algoritma K-Means," *J. Inform. dan Bisnis*, vol. 5, no. 1, pp. 13–29, 2016, [Online]. Available: <http://jurnal.kwikkiangie.ac.id/index.php/JIB/issue/viewIssue/41/69>.
- [5] T. H. Sardar and Z. Ansari, "Partition based clustering of large datasets using MapReduce framework: An analysis of recent themes and directions," *Futur. Comput. Informatics J.*, vol. 3, no. 2, pp. 247–261, 2018, doi: 10.1016/j.fcij.2018.06.002.

- [6] P. Anitha and M. M. Patil, "RFM model for customer purchase behavior using K-Means algorithm," *J. King Saud Univ. - Comput. Inf. Sci.*, no. xxxx, 2020, doi: 10.1016/j.jksuci.2019.12.011.
- [7] S. Learn, "Clustering," 2020. <https://scikit-learn.org/stable/modules/clustering.html> (accessed Nov. 27, 2020).
- [8] Y. Komaru, T. Yoshida, Y. Hamasaki, M. Nangaku, and K. Doi, "Hierarchical Clustering Analysis for Predicting 1-Year Mortality After Starting Hemodialysis," *Kidney Int. Reports*, vol. 5, no. 8, pp. 1188–1195, 2020, doi: 10.1016/j.ekir.2020.05.007.
- [9] T. Caliński and J. Harabasz, "A Dendrite Method For Cluster Analysis," *Commun. Stat.*, vol. 3, no. 1, pp. 1–27, 1974, doi: 10.1080/03610927408827101.
- [10] S. Asteriadis, K. Karpouzis, N. Shaker, and G. N. Yannakakis, "Towards detecting clusters of players using visual and gameplay behavioral cues," *Procedia Comput. Sci.*, vol. 15, pp. 140–147, 2012, doi: 10.1016/j.procs.2012.10.065.
- [11] D. L. Davies and D. W. Bouldin, "A Cluster Separation Measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, 1979, doi: 10.1109/TPAMI.1979.4766909.