

PERBANDINGAN METODE *K-MEANS* DAN *AVERAGE LINKAGE* UNTUK PENGELOMPOKKAN KABUPATEN/KOTA DI PROVINSI JAWA BARAT BERDASARKAN INDIKATOR KEMISKINAN

Salma Oktaviani^{1*}, Emeylia Safitri²

^{1,2}*Program Studi Statistika, Fakultas Sains dan Teknologi, Universitas Terbuka, Tangerang Selatan, Indonesia*

**Penulis korespondensi: salmaoktaviani33@gmail.com*

ABSTRAK

Kemiskinan merupakan salah satu permasalahan pembangunan yang masih menjadi perhatian di Provinsi Jawa Barat. Karakteristik kemiskinan yang berbeda antar kabupaten/kota menyebabkan dibutuhkan pengelompokan wilayah sebagai dasar penyusunan kebijakan agar lebih tepat sasaran. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Jawa Barat berdasarkan jumlah penduduk miskin, tingkat pengangguran terbuka (TPT), pengeluaran per kapita, gini ratio dan tingkat partisipasi angkatan kerja (TPAK) menggunakan metode *K-Means* dan *Average Linkage*. Penentuan jumlah *cluster* optimal pada metode *k-means* dilakukan menggunakan metode *silhouette*. Selanjutnya dilakukan pengelompokkan menggunakan metode *k-means* dan *average linkage* serta evaluasi kualitas *cluster* menggunakan indeks *silhouette*. Hasil penelitian menunjukkan bahwa jumlah *cluster* optimal yang diperoleh pada kedua metode adalah dua *cluster*. Nilai indeks *silhouette* pada metode *K-Means* sebesar 0,2831 dan pada metode *Average Linkage* sebesar 0,4631. Hasil tersebut menunjukkan bahwa metode *Average Linkage* menghasilkan kualitas pengelompokkan yang lebih baik dibandingkan metode *K-Means*. Oleh karena itu, metode *Average Linkage* dipilih sebagai metode terbaik dalam mengelompokkan kabupaten/kota di Provinsi Jawa Barat berdasarkan karakteristik kemiskinan.

Kata kunci: Kemiskinan, Analisis Cluster, *K-Means*, *Average Linkage*, *Silhouette*.

1 PENDAHULUAN

Kemiskinan adalah keadaan dimana seseorang memiliki pendapatan yang tidak cukup untuk memenuhi standar minimum hidup seperti makanan, pakaian, tempat tinggal dan kebutuhan sosial lainnya (Saputra et al., 2024). Faktor-faktor penyebab kemiskinan yaitu bisa karena kondisi alamiah dan ekonomi, kondisi struktural dan sosial, serta kondisi budaya. Masalah kemiskinan merupakan masalah yang kompleks dan bersifat multidimensional sehingga menjadi prioritas pembangunan (Ferezagia, 2018). Adapun upaya yang dilakukan oleh pemerintah setiap tahunnya untuk menyelesaikan masalah ini diantaranya yaitu bantuan sosial dan subsidi, pemberdayaan ekonomi & UMKM, pembangunan infrastruktur & akses dasar, pembaruan data & integrasi kebijakan, fokus khusus kemiskinan ekstrem (Assyafaah et al., 2025). Menurut data Badan Pusat Statistik (BPS) Persentase penduduk miskin pada September 2025 sebesar 8,25 persen, menurun 0,22 persen poin terhadap Maret 2025 dan menurun 0,32 persen poin terhadap September 2024 (Badan Pusat Statistik, 2026). Walaupun mengalami penurunan, namun angka tersebut masih tergolong cukup tinggi.

Ketimpangan pembangunan ekonomi antarwilayah menurut Sjafrizal (2012) merupakan fenomena umum yang terjadi dalam proses pembangunan ekonomi suatu daerah yang disebabkan oleh adanya perbedaan kandungan demografi yang terdapat pada masing-masing wilayah (Ginting, 2015). Akibatnya proses pertumbuhan ekonomi di setiap daerah menjadi berbeda-beda salah satu faktornya yaitu kondisi sumber daya alam. Perbedaan

kandungan sumber daya alam di setiap daerah jelas dapat mempengaruhi proses produksi pada daerah tersebut (Syafrizal, 2012). Daerah yang memiliki SDA yang melimpah mampu mempercepat proses produksi dan bisa dipasarkan dengan harga yang relatif murah, beda halnya dengan daerah yang memiliki SDA terbatas yang dimana akan berdampak dan proses produksi tersebut dan bisa dijual dengan harga yang tinggi sehingga sulit bersaing di pasar.

Salah satu analisis untuk mendeteksi ketimpangan tersebut adalah dengan analisis *cluster*. Analisis *Cluster* adalah metode dalam analisis data yang digunakan untuk mengelompokkan data berdasarkan tingkat kemiripan karakteristiknya dan yang berbeda akan dipisahkan (Talakua et al., 2017). Tujuan dari analisis *cluster* adalah mengelompokkan data yang memiliki karakteristik mirip ke dalam analisis *cluster* juga bertujuan untuk menemukan pola tersembunyi dalam *cluster* yang sama sehingga antar *cluster* memiliki perbedaan yang jelas. Analisis ini digunakan untuk menyederhanakan data yang kompleks menjadi beberapa kelompok yang lebih mudah dipahami dan dianalisis (Luthfi & Wijayanto, 2021). Selain itu, data, membantu pengambilan keputusan, serta menentukan prioritas kebijakan berdasarkan karakteristik masing-masing kelompok.

Penelitian sebelumnya telah membandingkan metode *hierarchical* dan *non-hierarchical*. Widyadhan et al. (2021) melakukan perbandingan metode *K-Means* dan *average linkage* pada data kemiskinan di Provinsi Jawa Tengah lalu diperoleh hasil bahwa metode *average linkage* menghasilkan *cluster* yang lebih baik dibandingkan metode *K-Means* dengan nilai *silhouette* yang lebih besar yaitu sebesar 0,35 (Widyadhan et al, 2021). Selain itu, terdapat penelitian yang dilakukan oleh Afira dan Wijayanto (2021) yang menggunakan metode *K-Means (partitioning)* dan *Ward (hierarki)* untuk mengelompokkan provinsi di Indonesia berdasarkan karakteristik kemiskinan yang hasilnya menyimpulkan bahwa metode *Ward* mampu menghasilkan pengelompokan terbaik. Agar lebih detail, penelitian ini hanya berfokus pada kemiskinan di Provinsi Jawa Barat dimana termasuk dalam provinsi memiliki tingkat kemiskinan yang cukup tinggi di Pulau Jawa. Penelitian ini akan menggunakan analisis *clustering* dengan metode *K-Means* dan *average linkage* untuk mengelompokkan tingkat kemiskinan kabupaten/kota di Provinsi Jawa Barat.

2. METODE

2.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Jawa Barat. Data yang digunakan dalam penelitian ini merupakan data tahun 2025. Adapun variabel yang digunakan dalam penelitian ini disajikan pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
Jumlah Penduduk Miskin	Menunjukkan berapa banyak penduduk yang berada di bawah garis kemiskinan.
TPT (Tingkat Pengangguran Terbuka)	Persentase jumlah pengangguran terhadap jumlah angkatan kerja. TPT yang tinggi menunjukkan lebih banyak orang tidak memiliki pendapatan dari pekerjaan.
Pengeluaran per kapita	Rata-rata pengeluaran yang dikeluarkan setiap penduduk dalam suatu wilayah selama periode tertentu. Semakin tinggi nilai

TPAK	pengeluarannya, biasanya semakin tinggi kemampuan masyarakat untuk memenuhi kebutuhan hidupnya. Persentase jumlah angkatan kerja dibandingkan dengan jumlah penduduk usia kerja. Angka ini mengindikasikan seberapa besar proporsi penduduk usia kerja (biasanya usia 15 tahun ke atas) yang aktif secara ekonomi, baik dengan bekerja maupun sedang mencari pekerjaan.
Gini Ratio	Menunjukkan tingkat ketimpangan distribusi pendapatan atau pengeluaran dalam suatu wilayah.

2.2 Analisis Cluster

Analisis *cluster* adalah metode yang digunakan untuk mengidentifikasi objek- objek yang homogen ke dalam kelompok-kelompok yang disebut *cluster* (Ramadani & Salma, 2022). *Cluster* yang terbentuk memiliki homogenitas internal yang tinggi dan heterogenitas eksternal yang tinggi (Hair, dkk., 2010:486). Analisis *clustering* umumnya dibagi menjadi 2 metode yaitu *hierarchical clustering* dan *non-hierarchical clustering* (Nafisah & Chandra, 2017). *Hierarchical clustering* dapat digunakan jika jumlah kelompok yang diinginkan tidak diketahui dan metode ini biasanya digunakan dalam objek dengan pengamatan yang tidak besar (Simamora, 2005). Sedangkan *non-hierarchical clustering* merupakan metode pengelompokan dimana *cluster* yang ingin dibentuk ditentukan terlebih dahulu, sehingga objek-objek akan dikelompokkan ke dalam kelompok yang telah ditentukan.

2.2.1 Hierarchical Clustering

Dalam metode hirarki *cluster* terdapat dua tipe dasar yaitu *agglomerative* (pemusatan) dan *divisive* (penyebaran). Dalam metode *agglomerative*, setiap obyek atau observasi dianggap sebagai sebuah *cluster* tersendiri. Dalam tahap selanjutnya, dua *cluster* yang mempunyai kemiripan digabungkan menjadi sebuah *cluster* baru demikian seterusnya. Selanjutnya, obyek atau observasi yang paling tinggi nilai ketidakmiripannya dipisahkan demikian seterusnya (B.J, 2009). Ada beberapa metode yang cukup terkenal dalam *agglomerative* (pemusatan) yaitu: *single linkage*, *complete linkage*, *average linkage*.

Metode *average linkage* adalah satu metode dalam *hierarchical clustering* (klaster hirarki) yang digunakan untuk mengelompokkan data berdasarkan rata-rata jarak antar anggota *cluster*. Jarak tersebut dapat dihitung dengan persamaan berikut (Suraya & Wijayanto, 2022):

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{(UV)}N_W} \quad (1)$$

Dimana d_{ik} merupakan jarak antar objek i dalam *cluster* (UV) dan objek k dalam *cluster* W . Sedangkan $N_{(UV)}$ dan N_W merupakan jumlah objek dalam *cluster* (UV) dan (W) (Johnson dan Wichern, 2007).

2.2.2 Non-Hierarchical Clustering

Non-hierarchical clustering biasa disebut juga sebagai *partitional clustering* merupakan metode pengelompokan data yang membagi data ke dalam sejumlah kelompok (klaster) yang telah ditentukan sejak awal (Sihombing et al., 2019). Metode yang paling umum digunakan dalam *non-hierarchical clustering* yaitu metode *K-means* yang dipopulerkan oleh

James McQueen. Metode ini dikembangkan agar analisis data dalam jumlah besar menjadi lebih mudah. Dalam prosesnya, data dikelompokkan berdasarkan kedekatan jarak nilai setiap titik. Keunggulan utama metode ini ada pada fleksibilitasnya. Berikut merupakan langkah-langkah algoritma *K-Means*, yaitu (Afidah & Masrukan, 2023):

1. Menentukan jumlah *cluster* (k) pada data *set* sebagai nilai *centroid*
2. Menghitung jarak antara data dan titik pusat *cluster* menggunakan rumus-rumus dari euclidian distance. Yang dapat dilihat pada teori jarak Euclidian yang dirumuskan sebagai berikut:

$$D_e = \sqrt{(X_i - S_i)^2 + (Y_i - t_i)^2} \quad (2)$$

Dimana:

D_e adalah *euclidean distance*

i adalah banyaknya objek

(x, y) adalah koordinat objek

(s, t) adalah koordinat *centroid*

3. Pusat *cluster* baru akan ditentukan bila semua data telah ditetapkan dalam *cluster* terdekat.
4. Proses penentuan titik pusat *cluster* dan penempatan data dalam *cluster* diulangi terus-menerus sampai nilai *centroid* tidak berubah lagi.

2.3 Tahapan Penelitian

Berikut tahapan- tahapan yang dilakukan untuk penelitian ini:

1. Mengumpulkan data penelitian.
2. Melakukan analisis statistika deskriptif untuk mengetahui gambaran umum karakteristik data penelitian.
3. Melakukan uji korelasi untuk mengetahui hubungan antar variabel.
4. Melakukan standarisasi data menggunakan metode *z-score* untuk menyamakan skala antar variabel.
5. Menentukan jumlah *cluster* optimal menggunakan metode *silhouette*
6. Melakukan pengelompokan data menggunakan metode *K-Means* berdasarkan jumlah *cluster* optimal yang diperoleh.
7. Melakukan pengelompokan data menggunakan metode *average linkage* berdasarkan jumlah *cluster* optimal yang diperoleh.
8. Menginterpretasikan dan membandingkan karakteristik masing-masing *cluster* yang terbentuk dari kedua metode.

3 HASIL DAN PEMBAHASAN

3.1 Statistika Deskriptif

Statistik deskriptif digunakan untuk menggambarkan karakteristik umum dari masing-masing variabel yang dianalisis (Saranty et al., 2025). Pada Tabel 2, disajikan statistika deskriptif terhadap 27 kabupaten / kota di Jawa Barat tahun 2025, diperoleh bahwa variabel Jumlah Penduduk Miskin memiliki rata rata sebesar 135,37 ribu jiwa dengan median sebesar 128,7 ribu jiwa. Lalu, variabel TPT memiliki rata rata sebesar 6,51% dengan median 6,6%, nilai maksimum 8,78% dan standar deviasi sebesar 1,6%, sehingga dapat dikatakan tingkat pengangguran antar daerah tidak terlalu berbeda jauh. Sementara variabel pengeluaran per kapita memiliki rata rata sebesar 1.636.086 rupiah dengan median 1.487.102 rupiah. Nilai maksimum mencapai 3.374.256 rupiah dan standar deviasi sebesar 553.405,6 yang menunjukkan adanya perbedaan tingkat pengeluaran masyarakat yang cukup besar antar kabupaten atau kota. Untuk variabel TPAK memiliki nilai mean 67,677 artinya dari 27 kabupaten/kota yang diamati setiap 100 penduduk usia kerja, sekitar 68 orang termasuk ke

dalam angkatan kerja. Untuk variabel gini ratio memiliki nilai rata-rata sebesar 0,365 dengan standar deviasi 0,039. Hal ini menunjukkan bahwa tingkat ketimpangan pendapatan di provinsi Jawa Barat relatif homogen dengan variansi yang tidak terlalu besar.

Tabel 2. Statistik Deskriptif

Variabel	Mean	Median	Maximum	Standar deviasi
Jumlah Penduduk Miskin (ribu)	135,37	128,7	401,9	89,8
TPT	6,51	6,6	8,78	1,61
Pengeluaran per kapita	1.636.086	1.487.102	3.374.256	553.405,6
TPAK	67,677	67,06	79,13	3,318
Gini Ratio	0,365	0,364	0,475	0,039

3.2 Uji Korelasi

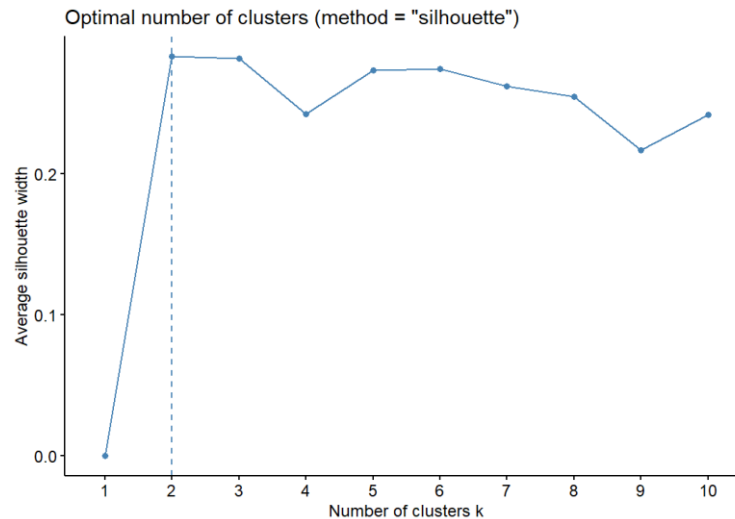
Uji korelasi dilakukan untuk melihat hubungan antar variabel sebelum dilakukan analisis *clustering*. Berdasarkan Tabel 3, terlihat bahwa nilai korelasi cenderung kecil hingga sedang. Korelasi positif yang kuat terdapat pada korelasi antara pengeluaran per kapita dan gini ratio dengan nilai 0,7806618.

Tabel 3. Hasil Uji Korelasi

	Jumlah Penduduk Miskin	TPT	Pengeluaran	Gini Ratio	TPAK
Jumlah Penduduk Miskin	1	0,1324938	-0,2691869	-0,5549818	-0,1522504
TPT	0,13249380	1	0,3592617	0,33199158	-0,6677177
Pengeluaran	-0,26918690	0,3592617	1	0,78066178	-0,5461164
Gini Ratio	-0,5549818	0,3319916	0,7806618	1	-0,5472916
TPAK	-0,15225035	-0,6677177	-0,5461164	-0,54729159	1

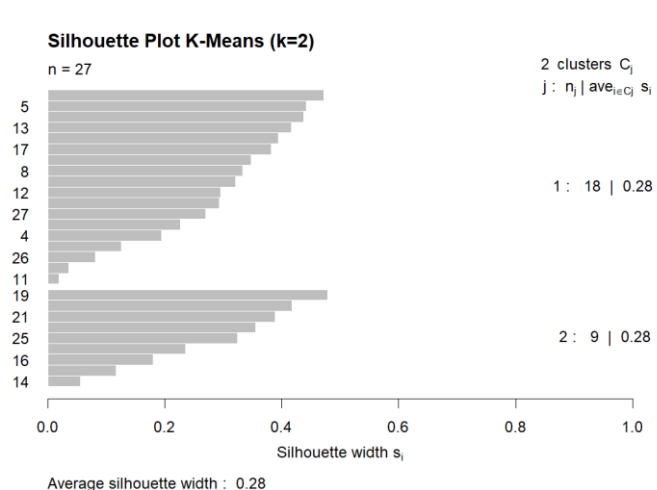
3.3 Hasil Metode K-Means

Sebelum melakukan analisis *clustering* perlu ditentukan dahulu jumlah *cluster*. Penentuan jumlah *cluster* dalam penelitian ini menggunakan metode *silhouette*. Dalam metode *silhouette* dapat diukur seberapa baik suatu objek diposisikan dalam sebuah *cluster*. Semakin tinggi nilai *silhouette width*, maka semakin baik kualitas *cluster* yang terbentuk.



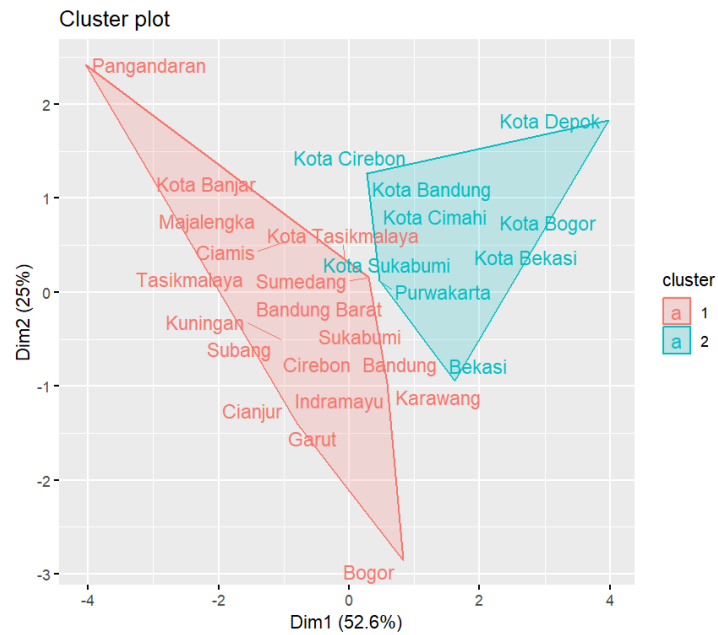
Gambar 1. *Silhouette* untuk Metode K-Means

Berdasarkan Gambar 1, nilai tertinggi diperoleh pada $k=2$. Setelah itu, semakin tinggi nilai k , nilai *silhouette* cenderung menurun. Hal tersebut menunjukkan bahwa pengelompokan data ke dalam dua *cluster* memberikan kualitas pemisahan yang paling baik dibandingkan jumlah *cluster* lainnya.



Gambar 2. *Silhouette Plot* Metode K- Means dengan $k=2$

Berdasarkan Gambar 2, diperoleh nilai rata-rata *silhouette* sebesar 0,28. Nilai tersebut menunjukkan bahwa hasil pengelompokan ke dalam 2 *cluster* memiliki kualitas yang cukup baik meskipun pemisahan antar cluster belum terlalu kuat. Selain itu, sebagian besar nilai *silhouette* berada pada sisi positif, yang mengindikasikan bahwa mayoritas objek telah terkelompok pada cluster yang sesuai.



Gambar 3. Cluster Plot Metode K-Means

Berdasarkan Tabel 4, Metode *K-Means* menghasilkan dua *cluster* dengan karakteristik berbeda. *Cluster 1* memiliki rata-rata jumlah penduduk miskin lebih tinggi sebesar 163,94 ribu jiwa dibanding *cluster 2* yaitu sebesar 78,23 ribu jiwa. Selain itu, *cluster 1* memiliki rata – rata pengeluaran perkapita lebih rendah yaitu Rp. 1.346.133 dibandingkan *cluster 2* Rp. 2.215.991. Hal tersebut menunjukkan bahwa tingkat kesejahteraan *cluster 2* lebih baik dibandingkan *cluster 1*.

Dari sisi ketenagakerjaan, *cluster 1* memiliki rata-rata nilai TPT sebesar 5,96% yang lebih rendah dibandingkan *cluster 2* yaitu 7,63%. Sementara itu nilai Tingkat Partisipasi Angkatan Kerja *cluster 1* memiliki nilai 68,64% lebih tinggi dibandingkan *cluster 2* yaitu 65,76%. Selain itu, nilai rata-rata Gini Ratio pada *cluster 1* sebesar 0,347 lebih rendah dibandingkan *cluster 2* sebesar 0,401 yang menunjukkan bahwa ketimpangan pendapatan pada *cluster 2* relatif lebih tinggi.

Tabel 4. Karakteristik Cluster Metode K-Means

Cluster	Penduduk Miskin	TPT	Pengeluaran	Gini Ratio	TPAK
1	163,94	5,96	1.346.133	0,347	68,64
2	78,23	7,63	2.215.991	0,401	65,76

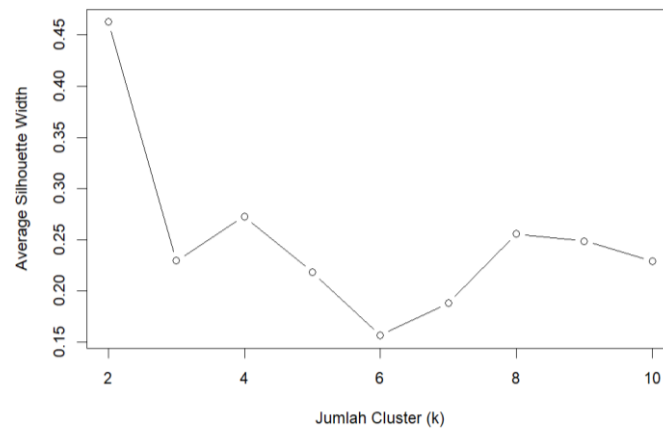
Tabel 5. Hasil Pengelompokkan Menggunakan Metode K-Means

Cluster	Wilayah
1	Kab. Bogor, Kab. Sukabumi, Kab. Cianjur, Kab. Bandung, Kab. Garut, Kab. Tasikmalaya, Kab. Ciamis, Kab. Kuningan, Kab. Cirebon, Kab. Majalengka, Kab. Sumedang, Kab. Indramayu, Kab. Subang, Kab. Karawang, Kab. Bandung Barat, Kab. Pangandaran, Kota Tasikmalaya, Kota Banjar.
2	Kab. Purwakarta, Kab. Bekasi, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi.

Berdasarkan Tabel 5, *cluster* 1 terdiri atas kabupaten/kota yang memiliki jumlah penduduk miskin relatif tinggi, TPT rendah, pengeluaran per kapita rendah dan tingkat partisipasi angkatan kerja tinggi. Sedangkan *cluster* 2 terdiri atas kabupaten/kota dengan tingkat pengeluaran per kapita yang lebih tinggi dan jumlah penduduk miskin yang relatif lebih rendah serta tingkat pengangguran terbuka dan ketimpangan pendapatan tinggi.

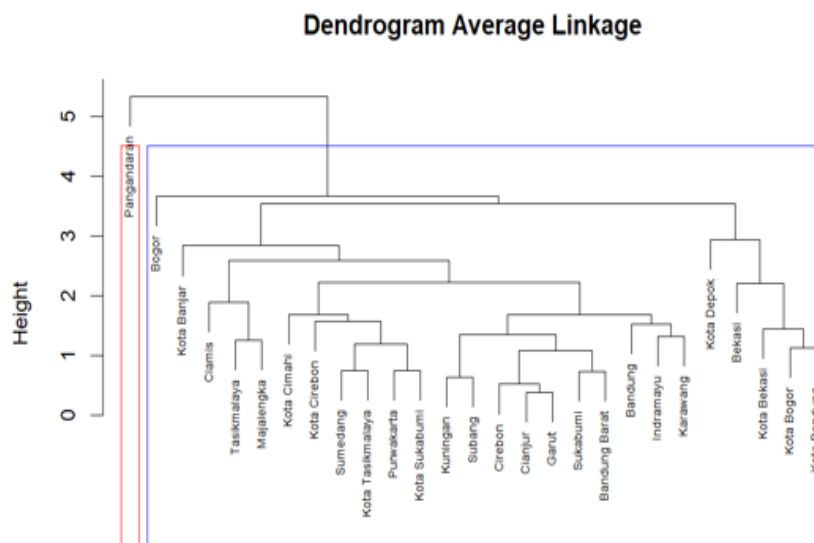
3.4 Metode Average Linkage

Dalam metode *average linkage*, matriks jarak diukur menggunakan jarak *euclidean* yang diperoleh dari data yang sudah di standarisasi.



Gambar 4. *Silhouette* untuk Metode *Average Linkage*

Berdasarkan Gambar 4, jumlah *cluster* optimal diperoleh pada $k=2$ karena menghasilkan nilai *silhouette* tertinggi sebesar sekitar 0,46. Nilai tersebut menunjukkan bahwa pengelompokan data ke dalam 2 cluster memberikan kualitas pemisahan cluster yang paling baik. Hasil pengelompokan objek berdasarkan rata-rata jarak tersebut divisualisasikan melalui dendrogram sebagai yang disajikan pada Gambar 5.



Gambar 5. Dendrogram Hasil Metode *Average Linkage*

Berdasarkan Tabel 6, dengan menggunakan metode *average linkage* terbentuk dua *cluster* yang karakternya cukup berbeda satu sama lain. *Cluster* 1 memiliki rata-rata jumlah penduduk miskin jauh lebih tinggi yaitu sekitar 139,3 ribu jiwa, sedangkan *cluster* 2 hanya 33,1 ribu jiwa. Lalu, dilihat dari pengeluaran per kapita, *cluster* 1 memiliki nilai yang lebih besar Rp1.645.912 dibanding *cluster* 2 yaitu Rp1.380.589. Tingkat Pengangguran Terbuka di *cluster* 1 juga lebih besar yaitu mencapai 6,69%, sedangkan *cluster* 2 sebesar 1,91%. Hal tersebut menunjukkan bahwa *cluster* 1 memiliki ketimpangan pendapatan yang lebih tinggi. Namun jika dilihat dari variabel Tingkat Partisipasi Angkatan Kerja *cluster* 1 memiliki nilai lebih rendah dibandingkan *cluster* 2. Selain itu, rata-rata gini ratio di *cluster* 1 tercatat 0,3671 lebih tinggi dibandingkan *cluster* 2 yang hanya 0,3040.

Tabel 6. Karakteristik *Cluster* Metode *Average Linkage*

<i>Cluster</i>	Penduduk Miskin	TPT	Pengeluaran	Gini Ratio	TPAK
1	139,304	6,693	1.645.912	0,3671	67,24
2	33,100	1,910	1.380.589	0,3040	79,13

Tabel 7. Hasil Pengelompokkan Menggunakan Metode *Average Linkage*

<i>Cluster</i>	Wilayah
1	Kab. Bogor, Kab. Sukabumi, Kab. Cianjur, Kab. Bandung, Kab. Garut, Kab. Tasikmalaya, Kab. Ciamis, Kab. Kuningan, Kab. Cirebon, Kab. Majalengka, Kab. Sumedang, Kab. Indramayu, Kab. Subang, Kab. Purwakarta, Kab. Karawang, Kab. Bekasi, Kab. Bandung Barat, Kota Bogor, Kota Bandung, Kota Bekasi, Kota Depok, Kota Sukabumi, Kota Cirebon, Kota Cimahi, Kota Tasikmalaya, Kota Banjar.
2	Kab. Pangandaran.

Berdasarkan Tabel 7, *Cluster* 1 terdiri atas kabupaten/kota dengan tingkat kemiskinan, pengangguran terbuka, dan ketimpangan pendapatan yang tinggi. Wilayah dalam *cluster* ini memerlukan perhatian yang lebih besar dalam upaya pengentasan kemiskinan dan peningkatan kesempatan kerja. Sedangkan *cluster* 2 terdiri atas daerah dengan kondisi sosial ekonomi yang relatif lebih baik. Hal ini ditunjukkan oleh jumlah penduduk miskin, tingkat pengangguran terbuka, dan ketimpangan pendapatan yang lebih rendah serta tingkat partisipasi angkatan kerja yang lebih tinggi. Dengan demikian Kab. Pangandaran dikategorikan sebagai wilayah dengan tingkat kesejahteraan yang lebih baik di Provinsi Jawa Barat berdasarkan variabel yang digunakan dalam penelitian ini.

Setelah dilakukan analisis *cluster* menggunakan metode *K-Means* dan *Average Linkage*, dilakukan perbandingan antara kedua metode yang disajikan pada Tabel 8.

Tabel 8. Hasil Perbandingan Metode *K-Means* dan *Average Linkage*

Metode	Nilai k Optimal	Nilai <i>Silhouette</i>
<i>K-Means</i>	2	0,2831
<i>Average Linkage</i>	2	0,4631

Berdasarkan hasil pengelompokan, dari kedua metode yang digunakan yaitu *K-Means* dan *Average Linkage* menghasilkan jumlah *cluster* optimal yang sama yaitu 2 *cluster* tetapi kualitas pengelompokan memiliki hasil yang berbeda yang ditunjukkan oleh nilai *silhouette coefficient*. Metode *K-Means* menghasilkan nilai *silhouette coefficient* sebesar 0,2831

sedangkan *Average Linkage* sebesar 0,4631. Nilai *silhouette* yang lebih tinggi menunjukkan bahwa objek dalam *cluster* memiliki tingkat kemiripan yang lebih tinggi dengan anggota cluster yang sama dan memiliki perbedaan yang lebih jelas terhadap cluster lainnya. Dengan demikian, metode *Average Linkage* memberikan hasil pengelompokan yang lebih baik dibandingkan metode *K-Means* pada data indikator kemiskinan kabupaten/kota di Provinsi Jawa Barat tahun 2025.

4 KESIMPULAN

Pada penelitian ini dilakukan perbandingan analisis *clustering* menggunakan metode *K-means* dan *average linkage* untuk mengetahui metode mana yang paling baik dalam pengelompokan kabupaten/kota berdasarkan indikator kemiskinan di Provinsi Jawa Barat. Dari hasil pengujian dengan nilai *silhouette coefficient*, kedua metode sama-sama menunjukkan bahwa jumlah cluster yang paling optimal adalah 2 *cluster*. Tetapi, metode *average linkage* mencatat nilai *silhouette* sebesar 0,4631 yang dimana lebih tinggi dari *K-Means* yaitu 0,2831. Hal ini menunjukkan bahwa metode *average linkage* lebih baik dan lebih akurat dalam mengelompokkan kabupaten/kota di Jawa Barat berdasarkan indikator kemiskinan. Sebaran anggota *cluster* menunjukkan Kab. Pangandaran membentuk *cluster* sendiri, sedangkan kabupaten/kota lainnya berada dalam *cluster* yang sama.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada seluruh pihak yang telah memberikan dukungan, bantuan, dan masukan selama proses pelaksanaan penelitian hingga penyusunan artikel ini.

DAFTAR PUSTAKA

- Afidah, N. N., & Masrukan. (2023). Penerapan metode clustering dengan algoritma K-means untuk pengelompokan data migrasi penduduk tiap kecamatan di Kabupaten Rembang. *PRISMA, Prosiding Seminar Nasional Matematika*, 6, 729–738. <https://journal.unnes.ac.id/sju/prisma/article/view/67038>
- Afira, N., & Wijayanto, A. W. (2021). Analisis cluster kemiskinan provinsi di Indonesia tahun 2019 dengan metode partitioning dan hierarki. *Komputika: Jurnal Sistem Komputer*, 10(2), 101–109. <https://doi.org/10.34010/komputika.v10i2.4317>
- Assyafaah, N., Dompok, T., & Salsabila, L. (2025). Upaya pemerintah dalam mengatasi tingkat kemiskinan di Indonesia. *Jurnal Publik: Jurnal Ilmiah Bidang Ilmu Administrasi Negara*, 19(2), 147–158. <https://doi.org/10.52434/jp.v19i02.444>
- Badan Pusat Statistik. (2026). *Persentase penduduk miskin September 2025 turun menjadi 8,25 persen*. <https://www.bps.go.id>
- Ferezagia, D. V. (2018). Analisis tingkat kemiskinan di Indonesia. *Jurnal Sosial Humaniora Terapan*, 1(1), 1–12. <https://doi.org/10.7454/jsht.v1i1.6>
- Ginting, A. M. (2015). Pengaruh ketimpangan pembangunan antarwilayah terhadap kemiskinan di Indonesia 2004–2013. *Kajian*, 20(1), 45–58.
- Luthfi, E., & Wijayanto, A. W. (2021). Analisis perbandingan metode hirearchical, K-Means, dan K-Medoids clustering dalam pengelompokan indeks pembangunan manusia Indonesia. *INOVASI: Jurnal Ekonomi, Keuangan, dan Manajemen*, 17(4), 761–773. <https://doi.org/10.30872/jinv.v17i4.10106>

- Nafisah, Q., & Chandra, N. E. (2017). Analisis cluster average linkage berdasarkan faktor-faktor kemiskinan di Provinsi Jawa Timur. *Zeta – Math Journal*, 3(2), 31–36. <https://doi.org/10.31102/zeta.2017.3.2.31-36>
- Ramadani, R., & Salma, A. (2022). Metode average linkage dan Ward dalam pengelompokan kesejahteraan Sumatera Barat tahun 2021. *Journal of Mathematics UNP*, 7(3), 11–24. <https://doi.org/10.24036/unpjomath.v7i3.12565>
- Saputra, D., Ardania, A., Putri, S., Asri, A. T. J. A., & Harsyiah, L. (2024). Analisis cluster untuk pengelompokan provinsi di Indonesia berdasarkan tingkat kemiskinan menggunakan metode average linkage. *Indonesian Journal of Applied Statistics and Data Science*, 1(1), 20–30. <https://doi.org/10.29303/ijasds.v1i1.5446>
- Saranty, Hidayatullah, M., & Fardinah. (2025). Analisis clustering berdasarkan faktor-faktor kemiskinan di Provinsi Sulawesi Selatan menggunakan algoritma K-Medoids. *Technologica*, 4(2), 181–194. <https://doi.org/10.55043/technologica.v4i2.342>
- Sihombing, R. E., Rachmatin, D., & Dahlan, J. A. (2019). *Program aplikasi bahasa R untuk pengelompokan objek menggunakan metode K-medoids clustering*. *Jurnal EurekaMatika*, 7(1), 71–82. <https://doi.org/10.17509/jem.v7i1.17888>
- Suraya, G. R., & Wijayanto, A. W. (2022). *Comparison of hierarchical clustering, K-means, K-medoids, and fuzzy C-means methods in grouping provinces in Indonesia according to the special index for handling stunting (Perbandingan metode hierarchical clustering, K-means, K-medoids, dan fuzzy C-means dalam pengelompokan provinsi di Indonesia menurut indeks khusus penanganan stunting)*. *Indonesian Journal of Statistics and Its Applications*, 6(2), 180–201. <https://doi.org/10.29244/ijsa.v6i2p180-201>
- Talakua, M. W., Leleury, Z. A., & Talluta, A. W. (2017). Analisis cluster dengan menggunakan metode K-Means untuk pengelompokan kabupaten/kota di Provinsi Maluku berdasarkan indikator indeks pembangunan manusia tahun 2014. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 11(2), 119–128. <https://doi.org/10.30598/barekengvol11iss2pp119-128>
- Widyadhan, D., Hastuti, R. B., Kharisudin, I., & Fauzi, F. (2021). Perbandingan analisis kluster K-Means dan average linkage untuk pengklasteran kemiskinan di Provinsi Jawa Tengah. *PRISMA, Prosiding Seminar Nasional Matematika*, 4, 584–594. <https://journal.unnes.ac.id/sju/index.php/prisma/article/view/45032>