

HIERARCHICAL AGGLOMERATIVE CLUSTERING DENGAN METODE WARD UNTUK PEMETAAN PASAR TENAGA KERJA PASCAPANDEMI DI JAWA TENGAH: PENDEKATAN *MACHINE LEARNING* BERBASIS KLASTERISASI

Fitri Rahmawati^{1,*}, Sefri Imanuel Fallo²

*¹Program Studi Matematika, Fakultas Pendidikan Matematika dan Ilmu
Pengetahuan Alam, Universitas Pendidikan Indonesia*

*²Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas San Pedro
email: fitrirahmawati@upi.edu*

Abstrak: Pandemi COVID-19 telah memberikan dampak signifikan terhadap dinamika ketenagakerjaan di Indonesia, termasuk di Provinsi Jawa Tengah. Ketimpangan distribusi pasar tenaga kerja antarwilayah menjadi tantangan tersendiri dalam perumusan kebijakan pascapandemi. Penelitian ini menerapkan Hierarchical Agglomerative Clustering (HAC) dengan metode Ward untuk memetakan pasar tenaga kerja pascapandemi di Jawa Tengah berdasarkan indikator Tingkat Partisipasi Angkatan Kerja (TPAK), Tingkat Pengangguran Terbuka (TPT), dan Rasio Pencari Kerja terhadap Lapangan Kerja (RPKL). Hasil klusterisasi menunjukkan lima klaster dengan karakteristik pasar tenaga kerja yang berbeda, mulai dari klaster dengan partisipasi kerja di bawah rata-rata dan pengangguran di atas rata-rata hingga klaster dengan kondisi pasar tenaga kerja yang lebih stabil. Validitas klaster dikonfirmasi melalui koefisien Silhouette. Temuan ini memberikan gambaran spasial yang berguna untuk perumusan kebijakan ketenagakerjaan yang adaptif dan berbasis data.

Kata Kunci: Hierarchical Agglomerative Clustering; pasar tenaga kerja; metode Ward

Abstract: The COVID-19 pandemic has had a significant impact on labor market dynamics in Indonesia, including in Central Java Province. The unequal distribution of labor market conditions across regions poses a specific challenge in formulating effective post-pandemic employment policies. This study applies Hierarchical Agglomerative Clustering (HAC) using Ward's method to map the post-pandemic labor market in Central Java based on Labor Force Participation Rate, Open Unemployment Rate, and Job Seeker to Job Vacancy Ratio. The clustering results identify five clusters with distinct labor market characteristics, ranging from clusters with low participation and high unemployment to clusters with more stable conditions. Cluster validity is confirmed by the Silhouette coefficient. These findings provide a spatial overview useful for formulating adaptive and data-driven labor policies.

Keywords: *Hierarchical Agglomerative Clustering; labor market; Ward's method*

PENDAHULUAN

Pandemi COVID-19 memberikan dampak sistemik terhadap pasar tenaga kerja di Indonesia, termasuk di Provinsi Jawa Tengah. Dampak tersebut tidak hanya tercermin dalam penurunan aktivitas ekonomi, tetapi juga secara signifikan mempengaruhi struktur ketenagakerjaan di tingkat regional. Data dari Badan Pusat Statistik (BPS) menunjukkan bahwa Tingkat Pengangguran Terbuka (TPT) di Jawa Tengah pada Agustus 2020 mencapai 6,48 persen, atau setara dengan 1,21 juta orang yang menganggur. Angka ini meningkat sebesar 2,04 persen poin dibandingkan Agustus 2019, yang berarti terdapat penambahan penganggur sebanyak 396 ribu orang (BPS Provinsi Jawa Tengah, 2020). Meskipun tren TPT mengalami penurunan secara bertahap setelah pandemi mereda, ketimpangan distribusi ketenagakerjaan antarwilayah di provinsi ini tetap menjadi tantangan utama dalam perumusan kebijakan pembangunan ketenagakerjaan yang adil dan berbasis data.

Analisis spasial berbasis statistik menjadi penting untuk memahami dinamika dan struktur pasar tenaga kerja lintas kabupaten/kota. Namun, pendekatan deskriptif sektoral yang selama ini digunakan belum mampu mengungkap pola kemiripan antarwilayah secara objektif dan terukur. Untuk itu, diperlukan metode analitik berbasis multivariat yang mampu mengelompokkan wilayah berdasarkan kemiripan karakteristik pasar tenaga kerja. Salah satu metode yang relevan untuk tujuan tersebut adalah *Hierarchical Agglomerative Clustering* (HAC), khususnya dengan pendekatan *Ward's Minimum Variance*. Secara global, perhatian terhadap dinamika pasar kerja juga terus meningkat seiring dengan transformasi struktural akibat perkembangan teknologi. Penelitian mengenai kecerdasan buatan (AI) dan pasar tenaga kerja berkembang pesat terutama pada topik partisipasi angkatan kerja, ketimpangan upah, stabilitas pekerjaan, dan adopsi teknologi AI dalam dunia kerja (Zhou et al., 2025).

Beberapa studi sebelumnya telah memanfaatkan metode klasterisasi dalam menganalisis indikator ketenagakerjaan untuk mengidentifikasi pola atau segmentasi wilayah berbasis data statistik. (Syafiyah et al., 2022) membandingkan efektivitas metode *hierarchical* dan *non-hierarchical clustering* dalam pengelompokan kabupaten/kota di Jawa Barat berdasarkan data ketenagakerjaan tahun 2020, dengan fokus pada pencarian struktur klaster yang paling representatif. Sementara itu, (Sebriana & Hasanah, 2025) menggunakan pendekatan clustering untuk mengevaluasi indikator ketenagakerjaan di Jawa Timur tahun 2023 dengan tujuan mengelompokkan wilayah menurut kemiripan kondisi pasar kerja. Kedua penelitian ini menunjukkan bahwa teknik klasterisasi dapat menjadi alat analisis statistik yang bermanfaat untuk mengeksplorasi karakteristik regional.

Hasil dari proses klusterisasi divisualisasikan dalam bentuk dendrogram, yang memperlihatkan hierarki penggabungan objek dan memungkinkan penentuan jumlah klaster yang optimal berdasarkan pola pemotongan cabang dendrogram (James et al., 2013). Dalam konteks ketenagakerjaan, indikator-indikator seperti Tingkat Pengangguran Terbuka (TPT), Tingkat Partisipasi Angkatan Kerja (TPAK), dan Rasio Penempatan terhadap Lowongan Kerja (RPLK) menjadi variabel penting yang mencerminkan kondisi dan efisiensi pasar tenaga kerja suatu wilayah. Dengan pendekatan ini, diharapkan dapat diperoleh segmentasi spasial wilayah kerja yang lebih informatif sebagai dasar perumusan kebijakan ketenagakerjaan berbasis data.

Meskipun beberapa studi sebelumnya telah menerapkan metode klusterisasi untuk menganalisis pasar tenaga kerja, kajian yang secara khusus mengimplementasikan HAC dalam konteks ketenagakerjaan pascapandemi di tingkat provinsi dengan pendekatan spasial masih terbatas. Terlebih lagi, penggunaan metrik evaluasi kuantitatif seperti koefisien Silhouette dalam menentukan jumlah klaster optimal belum banyak diadopsi dalam penelitian-penelitian sejenis. Padahal, pendekatan berbasis data sangat diperlukan dalam mendukung kebijakan publik yang responsif dan tepat sasaran. Dengan demikian, penelitian ini bertujuan untuk mengisi kekosongan tersebut melalui penerapan HAC berbasis indikator ketenagakerjaan yang divisualisasikan secara spasial, serta didukung oleh evaluasi objektif terhadap struktur klaster yang terbentuk.

METODE PENELITIAN

Penelitian ini menggunakan tiga variabel utama yang merepresentasikan kondisi pasar tenaga kerja di Provinsi Jawa Tengah pada tahun 2023, yaitu TPAK, TPT, dan Rasio Pencari Kerja terhadap Lapangan Kerja. Pemilihan ketiga variabel ini didasarkan pada relevansinya dalam menggambarkan keterlibatan penduduk usia kerja dalam aktivitas ekonomi, tingkat pengangguran yang terbuka di masyarakat, serta tekanan pasar tenaga kerja yang tercermin dari perbandingan antara jumlah pencari kerja dan kesempatan kerja yang tersedia. Sumber data yang digunakan dalam penelitian ini berasal dari publikasi resmi Badan Pusat Statistik (BPS) dan diakses melalui laman resmi BPS. Untuk memberikan pemahaman yang lebih jelas mengenai karakteristik masing-masing variabel yang digunakan, definisi operasional dari TPAK, TPT, dan Rasio Pencari Kerja terhadap Lapangan Kerja disajikan secara rinci pada tabel berikut:

Tabel 1. Deskripsi Variabel

Variabel	Keterangan
TPAK	Tingkat Partisipasi Angkatan Kerja adalah persentase jumlah angkatan kerja terhadap jumlah penduduk usia kerja (15 tahun ke atas).
TPT	Tingkat Pengangguran Terbuka adalah persentase jumlah pengangguran terhadap jumlah angkatan kerja.

Rasio Pencari Kerja terhadap Lapangan Kerja	RPKL merupakan perbandingan antara jumlah pencari kerja yang terdaftar dengan jumlah lowongan pekerjaan yang tersedia. Rasio ini digunakan untuk mengukur tingkat tekanan pasar kerja, yaitu seberapa besar jumlah pencari kerja dibandingkan dengan jumlah kesempatan kerja yang tersedia.
---	---

Klustering

Analisis kluster merupakan salah satu teknik dalam statistik multivariat yang berfungsi untuk mengelompokkan objek-objek ke dalam beberapa kelompok berdasarkan kemiripan karakteristik yang dimilikinya (Harnanto et al., 2017). Tujuan utama dari pendekatan ini adalah untuk menyusun objek yang serupa ke dalam satu kelompok sehingga pola-pola dalam data dapat dikenali dengan lebih jelas dan informatif. (Rahmawati et al., 2025) menyebutkan bahwa klustering merupakan metode statistik yang digunakan untuk mengelompokkan data ke dalam beberapa kluster berdasarkan kesamaan karakteristik. Dalam hal ini, kualitas hasil kluster ditentukan oleh tingkat homogenitas yang tinggi di dalam kluster (within-cluster) dan tingkat heterogenitas yang tinggi antar kluster (between-cluster), sebagaimana dijelaskan oleh (Imasdiani et al., 2022).

Metode klustering secara umum terbagi menjadi dua pendekatan utama, yaitu klustering hirarkis dan non-hirarkis (Apriliana & Widodo, 2023). Klustering hirarkis, seperti metode agglomerative dan divisive, membentuk struktur pohon (dendrogram) yang merepresentasikan hubungan antar objek berdasarkan kedekatan atau kemiripan tertentu. Sebaliknya, klustering non-hirarkis, seperti K-Means dan K-Medoids, mengelompokkan data ke dalam jumlah kluster yang telah ditentukan sebelumnya tanpa membentuk struktur hierarki.

Hierarchical Agglomerative Clustering dengan Metode Ward

Hierarchical Agglomerative Clustering (HAC) merupakan salah satu teknik klasterisasi hirarkis yang banyak digunakan dalam analisis data multivariat. HAC bekerja dengan cara menggabungkan objek-objek yang paling mirip secara iteratif, dimulai dari setiap objek sebagai kluster tunggal, kemudian menggabungkannya secara berurutan hingga terbentuk satu kluster besar yang mencakup seluruh objek (Li et al., 2022). Menurut (James et al., 2013), salah satu karakteristik utama dari pendekatan ini adalah visualisasi hasil pengelompokan dalam bentuk dendrogram, yang menggambarkan hubungan antar objek dan kluster secara hirarkis. Hasil proses ini divisualisasikan dalam dendrogram yang memperlihatkan struktur pengelompokan secara bertingkat dan memungkinkan penentuan jumlah kluster yang optimal berdasarkan pola pemotongan dendrogram pada ketinggian tertentu. Meskipun posisi horizontal antar objek pada dendrogram tidak dapat digunakan untuk menyimpulkan

kemiripan, posisi vertikal titik penggabungan mencerminkan tingkat kesamaan, sehingga memberikan informasi penting untuk interpretasi kluster secara tepat.

Metode Ward dalam HAC adalah pendekatan penggabungan kluster yang bertujuan meminimalkan peningkatan variansi total dalam kluster pada setiap tahap penggabungan. Secara matematis, Ward meminimalkan jumlah kuadrat jarak antar anggota kluster terhadap pusat kluster sehingga menghasilkan kluster yang relatif homogen dan seimbang. Proses komputasi metode Ward dimulai dengan menghitung matriks jarak antar objek, lalu secara iteratif menggabungkan dua kluster yang menghasilkan peningkatan variansi terkecil (Eszergár-Kiss & Caesar, 2017). Menurut (Murtagh & Legendre, 2014), metode Ward menggunakan jarak kuadrat antar titik (squared Euclidean distance) secara konsisten sepanjang proses penggabungan. Ini menjadikannya metode yang secara matematis lebih stabil.

Langkah-langkah Komputasi HAC dengan Ward:

1. Menghitung jarak kuadrat antar setiap pasang objek.

Langkah pertama adalah menghitung jarak antar semua pasangan objek menggunakan *squared Euclidean distance*, yaitu:

$$d^2(x_i, x_j) = \sum_{k=1}^p (x_{ik} - x_{jk})^2$$

dengan:

$d^2(x_i, x_j)$ = jarak kuadrat eucildean objek i dan j,

p = jumlah variabel atau dimensi,

x_{ik} = nilai variabel ke-k pada objek i.

2. Mengisialisasi kluster tunggal.

Setiap objek awalnya dianggap sebagai kluster tersendiri, sehingga terdapat n kluster tunggal (singleton clusters).

3. Identifikasi pasangan kluster.

Menentukan pasangan kluster yang jika digabung akan menghasilkan kenaikan error sum of squares (ESS) paling kecil.

4. Memperbarui jarak antara kluster.

Setelah dua kluster A dan B digabungkan menjadi $A \cup B$, maka jarak kuadrat antar kluster diperbarui terhadap kluster lain C menggunakan formula:

$$D_{A \cup B, C} = \frac{n_A + n_C}{n_A + n_B + n_C} d^2(\bar{x}_A, \bar{x}_C) + \frac{n_B + n_C}{n_A + n_B + n_C} d^2(\bar{x}_B, \bar{x}_C) - \frac{n_C}{n_A + n_B + n_C} d^2(\bar{x}_A, \bar{x}_B)$$

dengan:

$D_{A \cup B, C}$ = jarak kuadrat gabungan kluster $A \cup B$ dengan kluster C,

n_A, n_B, n_C = jumlah anggota masing-masing kluster A, B, C,

$\bar{x}_A, \bar{x}_B, \bar{x}_C$ = rata-rata (centroid) masing-masing kluster A, B, C,

5. Mengulangi langkah 3 dan 4 sampai seluruh objek tergabung dalam satu kluster tunggal.

-
6. Hasil akhir divisualisasikan dalam bentuk dendrogram, yang memperlihatkan hierarki penggabungan klaster.

Koefisien Silhouette

Koefisien Silhouette merupakan salah satu metode evaluasi internal yang banyak digunakan untuk menilai kualitas hasil klustering. Metode ini mengukur sejauh mana sebuah objek berada dalam klaster yang tepat, dengan membandingkan kedekatannya terhadap anggota dalam klasternya sendiri dengan kedekatannya terhadap anggota klaster lain (Ishaifa, 2025). Nilai koefisien Silhouette dirumuskan sebagai:

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

dengan:

$a(i)$ = Rata-rata jarak antara objek ke- i dan seluruh objek lain dalam klaster yang sama

$b(i)$ = nilai minimum dari rata-rata jarak antara objek ke- i dan seluruh objek dalam klaster yang berbeda.

Rahmawati (2025) menyatakan Koefisien Silhouette memiliki nilai yang berkisar antara -1 hingga 1. Semakin tinggi nilainya mendekati 1, menunjukkan bahwa objek semakin sesuai dengan klaster tempatnya berada. Sebaliknya, nilai yang mendekati -1 mengindikasikan pengelompokan objek kurang tepat.

HASIL DAN PEMBAHASAN

Hasil Penelitian

Analisis Deskriptif

Analisa ini menyajikan deskripsi data untuk memberikan gambaran awal mengenai kondisi pasar tenaga kerja di Provinsi Jawa Tengah tahun 2023 berdasarkan tiga variabel utama, yaitu TPAK, TPT, dan RPKL. Ketiga variabel ini merefleksikan partisipasi, pengangguran, dan tekanan dalam pasar kerja yang menjadi dasar analisis klaster selanjutnya.

Tabel 2. Analisis Deskriptif

Variabel	Rata-rata	Median	Nilai Minimum	Nilai Maksimum
TPAK	72.11	72.09	64.60	82.45
TPT	4.839	4.570	1.920	8.980
RPKL	2.2692	1.7266	0.1668	12.2519

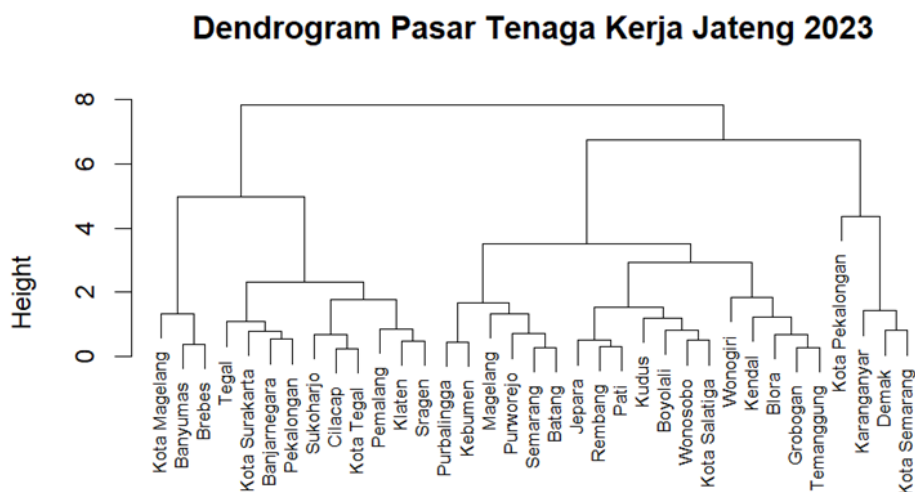
Tabel 2 menunjukkan analisis deskriptif terhadap variabel-variabel pasar tenaga kerja di Provinsi Jawa Tengah tahun 2023. Pada variabel Tingkat Partisipasi Angkatan Kerja (TPAK), nilai minimum tercatat di Kabupaten Banyumas sebesar 64,60%, sementara nilai maksimum terdapat di Kabupaten Wonogiri sebesar 82,40%. Rata-rata

TPAK secara keseluruhan adalah 72,11%. Nilai-nilai ini mencerminkan perbedaan tingkat keterlibatan penduduk usia kerja dalam aktivitas ekonomi di masing-masing daerah.

Pada variabel Tingkat Pengangguran Terbuka (TPT), nilai terendah ditemukan di Kabupaten Karanganyar sebesar 1,92%, sedangkan nilai tertinggi tercatat di Kota Magelang sebesar 8,98%, dengan rata-rata sebesar 4,839%. Rentang nilai ini menunjukkan adanya variasi dalam proporsi pengangguran terbuka di antara kabupaten/kota di Jawa Tengah. Pada variabel Rasio Pencari Kerja terhadap Lapangan Kerja (RPKL), nilai minimum tercatat di Kabupaten Sragen sebesar 0,166, dan nilai maksimum terdapat di Kota Pekalongan sebesar 12,25, dengan rata-rata provinsi sebesar 2,2692. Nilai RPKL ini menggambarkan perbedaan jumlah pencari kerja relatif terhadap jumlah lowongan kerja yang tersedia di masing-masing daerah. Nilai-nilai tersebut memberikan gambaran awal mengenai distribusi dan karakteristik pasar tenaga kerja antar wilayah di Provinsi Jawa Tengah, yang selanjutnya akan menjadi dasar dalam analisis kluster untuk mengelompokkan daerah berdasarkan kemiripan karakteristik pasar kerjanya.

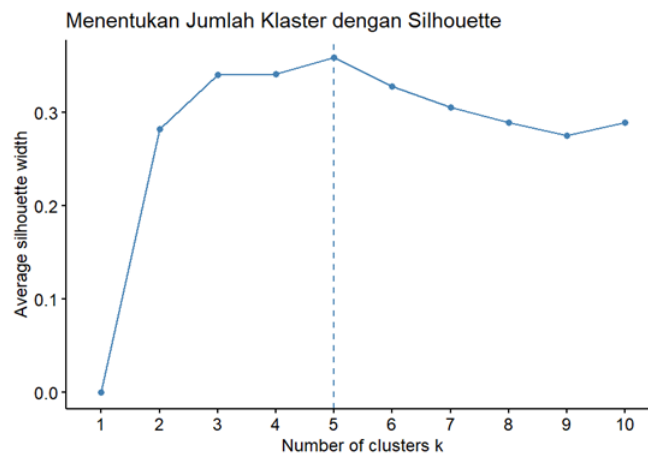
Analisis Clustering

Setelah dilakukan analisis deskriptif terhadap masing-masing variabel, tahap selanjutnya adalah melakukan analisis kluster untuk mengelompokkan kabupaten/kota di Provinsi Jawa Tengah berdasarkan karakteristik pasar tenaga kerja. Teknik klasterisasi yang digunakan dalam penelitian ini adalah Hierarchical Agglomerative Clustering (HAC) dengan metode Ward. Data terlebih dahulu dilakukan standarisasi. Hasil dari proses ini divisualisasikan dalam bentuk dendrogram, yang memberikan representasi struktur hierarki pengelompokan antar wilayah berdasarkan kemiripan nilai pada variabel TPAK, TPT, dan RPKL.



Gambar 1. Dendrogram

Berdasarkan dendrogram pada Gambar 1, langkah berikutnya adalah menentukan jumlah kluster yang paling sesuai. Untuk memperoleh dasar yang lebih objektif dalam penentuan jumlah kluster, dilakukan evaluasi menggunakan *silhouette score*. Metode ini mengukur seberapa baik setiap objek cocok dengan klasternya dibandingkan dengan kluster lain, sehingga membantu menilai kualitas pemisahan kluster yang terbentuk.

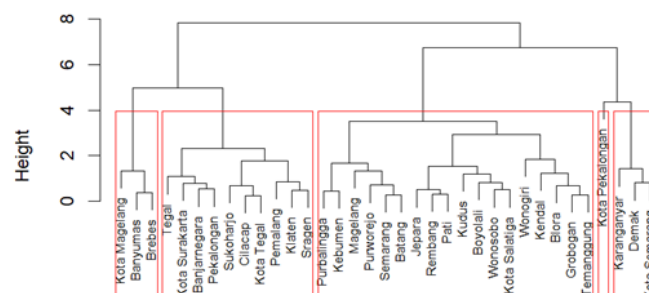


Gambar 2. Grafik Silhouette Score

Berdasarkan hasil evaluasi menggunakan grafik silhouette pada Gambar 2, diperoleh bahwa jumlah kluster terbaik untuk pengelompokan wilayah di Provinsi Jawa Tengah adalah sebanyak lima kluster ($k = 5$). Nilai rata-rata silhouette tertinggi pada $k = 5$ menunjukkan bahwa pembagian data ke dalam lima kelompok memberikan pemisahan kluster yang lebih baik dibandingkan pilihan jumlah kluster lainnya. Dengan demikian, hasil ini menjadi dasar yang kuat untuk menentukan jumlah kluster optimal dalam analisis selanjutnya.

Setelah dilakukan evaluasi jumlah kluster menggunakan metode silhouette yang menunjukkan bahwa lima kluster merupakan jumlah yang optimal, tahap selanjutnya adalah memvisualisasikan hasil pengelompokan tersebut dengan dendrogram. Dendrogram dengan pemotongan pada lima kluster ini akan memperlihatkan struktur hierarki pengelompokan serta distribusi wilayah berdasarkan karakteristik variabel yang digunakan.

Dendrogram Pasar Tenaga Kerja Jateng 2023 dengan 5 Kluster



Gambar 3. Dendrogram dengan lima kluster

Gambar 3 menampilkan dendrogram dengan pemotongan pada lima klaster. Dendrogram ini menggambarkan pembagian wilayah di Provinsi Jawa Tengah ke dalam kelompok-kelompok yang memiliki karakteristik serupa berdasarkan variabel yang dianalisis. Untuk memudahkan interpretasi dan analisis lebih lanjut, Tabel 3 berikut menyajikan daftar kota dan kabupaten yang tergolong ke dalam masing-masing klaster sesuai hasil pengelompokan.

Tabel 3. Klasterisasi Kabupaten/Kota

Klaster	Kabupaten/Kota
1	Tegal, Kota Surakarta, Banjarnegara, Pekalongan, Sukoharjo, Cilacap, Kota Tegal, Pemalang, Klaten, Sragen
2	Kota Magelang, Banyumas, Brebes
3	Purbalingga, Kebumen, Magelang, Purworejo, Semarang, Batang, Jepara, Rembang, Pati, Kudus, Boyolali, Wonosobo, Kota Salatiga, Wonogiri, Kendal, Blora, Grobogan, Temanggung
4	Kota Semarang, Demak, Karanganyar
5	Kota Pekalongan

Pembahasan

Setelah mengelompokkan kota dan kabupaten ke dalam lima klaster, langkah berikutnya adalah menganalisis karakteristik masing-masing klaster melalui pusat klaster (centroid) dan standar deviasi. Centroid ini merepresentasikan nilai rata-rata variabel (Raja et al., 2024), sehingga memberikan gambaran profil khas dari kelompok wilayah tersebut berdasarkan variabel Tingkat Partisipasi Angkatan Kerja (TPAK), Tingkat Pengangguran Terbuka (TPT), dan Rasio Pencari Kerja terhadap Lapangan Kerja (RPKL).

Tabel 4. Centroid dan Standar Deviasi Tiap Klaster

Klaster	Jumlah Kabupaten/Kota	Rata-rata TPAK	Standar Deviasi TPAK	Rata-rata TPT	Standar Deviasi TPT	Rata-rata RPKL	Standar Deviasi RPKL
1	10	-0.881	0.420	0.0646	0.595	-0.411	0.377
2	3	-1.29	0.629	2.35	0.115	-0.474	0.256
3	18	0.753	0.679	-0.310	0.694	-0,192	0.372
4	3	-0,44	0.183	-0,929	0.791	1,48	0.0941
5	1	0.446		0.686		4,54	

Berdasarkan Tabel 4, nilai standar deviasi antar wilayah dalam masing-masing klaster berkisar antara 0,094 hingga 0,79. Rentang tersebut menunjukkan bahwa keragaman internal dalam tiap klaster masih berada dalam batas yang dapat diterima. Meskipun terdapat variasi antar wilayah dalam satu klaster, karakteristik umumnya tetap

cukup konsisten, sehingga klaster yang terbentuk dapat dianggap representatif untuk dianalisis sebagai satu kelompok.

Mengingat data telah distandarisasi, data akan memiliki karakteristik nilai rata-rata sebesar 0 dan simpangan baku sebesar 1 (Talakua et al., 2017). Dengan demikian nilai negatif pada centroid mengindikasikan bahwa nilai tersebut berada di bawah rata-rata, adapun nilai positif mengindikasikan centroid di atas rata-rata (Pramudita & Sumargo, 2019). Klaster 1 dan Klaster 2 menunjukkan karakteristik yang serupa, yaitu nilai Tingkat Partisipasi Angkatan Kerja (TPAK) dan Rasio Pencari Kerja terhadap Lapangan Kerja (RPKL) yang berada di bawah rata-rata, sementara Tingkat Pengangguran Terbuka (TPT) berada pada level yang tinggi. Kondisi ini mengindikasikan bahwa pada wilayah-wilayah dalam klaster ini partisipasi angkatan kerja relatif rendah, jumlah pencari kerja terhadap lapangan kerja tersedia juga sedikit, namun tingkat pengangguran terbuka cukup tinggi. Situasi ini dapat mencerminkan tantangan dalam penyerapan tenaga kerja serta potensi ketidaksesuaian antara kebutuhan pasar kerja dengan ketersediaan lapangan kerja di wilayah tersebut.

Klaster 3 memiliki ciri khas dengan nilai TPAK yang tinggi, sedangkan TPT dan RPKL berada di bawah rata-rata. Hal ini menunjukkan bahwa wilayah dalam klaster ini memiliki partisipasi angkatan kerja yang tinggi dan tingkat pengangguran serta rasio pencari kerja terhadap lapangan kerja yang relatif rendah. Profil ini mengindikasikan kondisi pasar tenaga kerja yang lebih stabil dan sehat, dengan tingkat keterlibatan angkatan kerja yang tinggi serta tekanan pengangguran dan persaingan pencari kerja yang rendah.

Klaster 4 memperlihatkan nilai TPAK dan TPT di bawah rata-rata, namun Rasio Pencari Kerja terhadap Lapangan Kerja (RPKL) justru di atas rata-rata. Artinya, meskipun partisipasi angkatan kerja dan tingkat pengangguran terbuka relatif rendah, jumlah pencari kerja yang tersedia dibandingkan dengan lapangan kerja masih cukup besar. Hal ini bisa mengindikasikan potensi tekanan pada pasar tenaga kerja meskipun tingkat pengangguran saat ini belum menunjukkan angka yang tinggi.

Klaster 5 menunjukkan nilai di atas rata-rata pada ketiga variabel, yaitu TPAK, TPT, dan RPKL. Dengan kata lain, wilayah-wilayah dalam klaster ini memiliki tingkat partisipasi angkatan kerja yang relatif tinggi, tingkat pengangguran terbuka yang juga tinggi, serta rasio pencari kerja terhadap lapangan kerja yang besar. Kondisi ini menggambarkan dinamika pasar tenaga kerja yang kompleks, di mana meskipun banyak angkatan kerja yang aktif, tekanan pengangguran dan persaingan untuk memperoleh pekerjaan tetap tinggi.

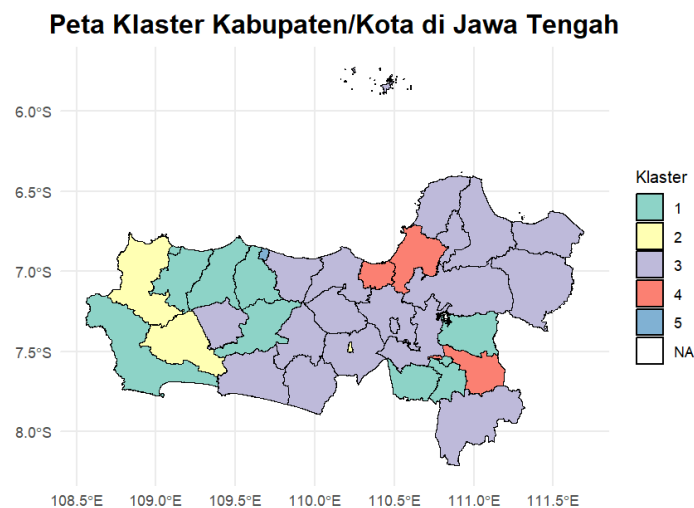
Dari kelima klaster tersebut, perhatian khusus perlu diberikan pada Klaster 1, Klaster 2, Klaster 4, dan Klaster 5. Klaster 1 dan Klaster 2 memiliki tingkat partisipasi angkatan kerja (TPAK) dan rasio pencari kerja terhadap lapangan kerja (RPKL) yang di bawah rata-rata, namun diiringi dengan tingkat pengangguran terbuka (TPT) yang di

atas rata-rata. Kondisi ini menunjukkan adanya tantangan dalam penyerapan tenaga kerja dan keterbatasan peluang kerja, sehingga intervensi kebijakan yang fokus pada peningkatan kesempatan kerja dan pengembangan keterampilan sangat diperlukan di wilayah tersebut.

Klaster 4 menunjukkan nilai TPAK dan TPT yang rendah, namun rasio pencari kerja terhadap lapangan kerja (RPKL) tinggi, menandakan adanya potensi tekanan pada pasar tenaga kerja yang belum tercermin pada tingkat pengangguran saat ini. Oleh karena itu, klaster ini juga perlu mendapatkan perhatian agar tekanan tersebut tidak berkembang menjadi permasalahan yang lebih serius di masa mendatang.

Sementara itu, Klaster 5 meskipun memiliki partisipasi angkatan kerja yang tinggi, juga menghadapi tingkat pengangguran dan rasio pencari kerja terhadap lapangan kerja yang tinggi. Hal ini menunjukkan adanya persaingan yang ketat di pasar tenaga kerja dan membutuhkan strategi yang efektif dalam penyerapan tenaga kerja serta penyesuaian keterampilan agar sesuai dengan kebutuhan pasar. Sebaliknya, Klaster 3 menunjukkan kondisi pasar tenaga kerja yang relatif lebih stabil dan sehat, sehingga prioritas intervensi dapat lebih difokuskan pada klaster-klaster yang menunjukkan potensi masalah seperti Klaster 1, 2, 4, dan 5.

Visualisasi peta klaster dapat dilihat pada Gambar 4 berikut.



SIMPULAN

Penelitian ini menggunakan metode *Hierarchical Agglomerative Clustering* (HAC) dengan pendekatan Ward untuk memetakan kondisi pasar tenaga kerja pascapandemi di Provinsi Jawa Tengah berdasarkan indikator Tingkat Partisipasi Angkatan Kerja (TPAK), Tingkat Pengangguran Terbuka (TPT), dan Rasio Pencari Kerja terhadap Lapangan Kerja (RPKL). Evaluasi menggunakan koefisien Silhouette

menunjukkan bahwa pemilihan jumlah klaster optimal adalah lima. Hasil klasterisasi menunjukkan heterogenitas kondisi pasar tenaga kerja antarwilayah. Klaster 1 dan 2 ditandai dengan partisipasi kerja dan RPKL rendah namun TPT tinggi, yang mencerminkan tantangan dalam penyerapan tenaga kerja. Klaster 3 mencerminkan kondisi ideal dengan TPAK tinggi dan tekanan pengangguran rendah. Sementara Klaster 4 dan 5 menggambarkan tekanan pasar kerja melalui tingginya RPKL, meskipun partisipasi angkatan kerja bervariasi. Dengan demikian, pendekatan HAC yang dievaluasi secara objektif dengan koefisien Silhouette memberikan pemetaan spasial yang informatif, yang dapat digunakan sebagai dasar penyusunan kebijakan ketenagakerjaan yang adaptif dan berbasis bukti.

DAFTAR PUSTAKA

- Apriliansa, T., & Widodo, E. (2023). Analisis Cluster Hierarki untuk Pengelompokan Provinsi di Indonesia berdasarkan Jumlah Base Transceiver Station dan Kekuatan Sinyal. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 3(2), 286–296. <https://doi.org/10.24002/konstelasi.v3i2.7143>
- BPS Provinsi Jawa Tengah. (2020). Keadaan Ketenagakerjaan Provinsi Jawa Tengah, Agustus 2020: Tingkat Pengangguran Terbuka (TPT) sebesar 6,48 persen, lebih tinggi dari Agustus 2019 yang sebesar 4,44 persen. In *Berita Resmi Statistik*.
- Eszergár-Kiss, D., & Caesar, B. (2017). Definition of user groups applying Ward's method. *Transportation Research Procedia*, 22(2017), 25–34. <https://doi.org/10.1016/j.trpro.2017.03.004>
- Harnanto, Y. I., Rusgiyono, A., & Wuryandari, T. (2017). Penerapan Analisis Klaster Metode Ward Terhadap Kabupaten/Kota Di Jawa Tengah Berdasarkan Pengguna Alat Kontrasepsi. *Jurnal Gaussian*, 6(4), 528–537. <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- Imasdiani, I., Purnamasari, I., & Amijaya, F. D. T. (2022). Perbandingan Hasil Analisis Cluster Dengan Menggunakan Metode Average Linkage Dan Metode Ward. *Eksponensial*, 13(1), 9–17. <https://doi.org/10.30872/eksponensial.v13i1.875>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning - with Applications in R. In *Book*. Springer.
- Li, T., Rezaeipanah, A., & Tag El Din, E. S. M. (2022). An ensemble agglomerative hierarchical clustering algorithm based on clusters clustering technique and the novel similarity measurement. *Journal of King Saud University - Computer and Information Sciences*, 34(6), 3828–3842. <https://doi.org/10.1016/j.jksuci.2022.04.010>
- Murtagh, F., & Legendre, P. (2014). Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *Journal of Classification*, 31, 275–295. <https://doi.org/10.1128/jb.160.1.233-238.1984>
- Pramudita, D. A., & Sumargo, B. (2019). Pengelompokan Pengguna Internet dengan Metode K-Means Clustering. *Jurnal Statistika Dan Aplikasinya*, 3(1), 1–12. <https://doi.org/10.21009/jsa.03101>
- Rahmawati, F., Putri, N. S., Ulfa, N., & Azizah, U. N. (2025). KLASTERISASI

KABUPATEN/KOTA DI JAWA BARAT BERDASARKAN KUALIFIKASI DAN PEMERATAAN GURU SMP MENGGUNAKAN K-MEANS CLUSTERING. *ProSandika*, 6, 149–158.

Raja, N. A., Tinungki, G. M., & Sirajang, N. (2024). Implementasi Algoritma Centroid Linkage dan K-Medoids dalam Mengelompokkan Kabupaten/Kota di Sulawesi Selatan Berdasarkan Indikator Pendidikan. *ESTIMASI: Journal of Statistics and Its Application*, 5(1), 61–74. <https://doi.org/10.20956/ejsa.v5i1.13605>

Sebriana, E. I., & Hasanah, S. H. (2025). Analisis Indikator Ketenagakerjaan di Jawa Timur Tahun 2023 dengan Pendekatan Clustering. *Prosiding Seminar Nasional Sains Dan Teknologi Seri III*, 2(1), 1056–1068.

Syafiyah, U., Asrafi, I., Wicaksono, B., Puspitasari, D. P., & Sirait, M. (2022). Analisis Perbandingan Metode Cluster Data Indikator Ketenagakerjaan di Jabar Tahun 2020. *Seminar Nasional Official Statistics*, 803–812. <https://doi.org/10.34123/semnasoffstat.v2022i1.1221>

Talakua, M. W., Leleury, Z. A., & Taluta, A. W. (2017). Analisis Cluster Dengan Menggunakan Metode K-Means Untuk Pengelompokkan 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>

Zhou, H., Wang, L., Cao, Y., & Li, J. (2025). The impact of artificial intelligence on labor market: A study based on bibliometric analysis. *Journal of Asian Economics*, 98(January), 101926. <https://doi.org/10.1016/j.asieco.2025.101926>