

## ANALISIS INDIKATOR KETENAGAKERJAAN DI JAWA TIMUR TAHUN 2023 DENGAN PENDEKATAN *CLUSTERING*

Eva Intan Sebriana\*, Siti Hadijah Hasanah

Program Studi Statistika, Universitas Terbuka, Tangerang Selatan

\*Penulis korepondensi: [evansebriana@gmail.com](mailto:evansebriana@gmail.com)

### ABSTRAK

Ketenagakerjaan di Jawa Timur memiliki dinamika yang mencerminkan variasi kondisi antarwilayah, baik di daerah perkotaan maupun pedesaan. Ketidakseimbangan akses terhadap lapangan kerja, pendidikan, dan sektor ekonomi menjadi tantangan utama yang memengaruhi distribusi tenaga kerja. Perubahan struktur ekonomi dari sektor agraris ke sektor industri dan jasa juga turut memengaruhi pola ketenagakerjaan di berbagai wilayah. Pendekatan clustering dapat digunakan untuk mengelompokkan wilayah kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator ketenagakerjaan yang serupa. Analisis yang akan digunakan yaitu metode hierarchical clustering dan non-hierarchical clustering. Hasil analisis menunjukkan bahwa metode *clustering* paling baik adalah metode *hierarchical clustering (complete linkage)*. Pemilihan metode *clustering* terbaik yaitu dengan melihat nilai validitas metode *connectivity* terkecil serta *silhouette* dan *dunn* terbesar. Diperoleh jumlah cluster sebanyak 2 *cluster* dimana *cluster 1* terdiri atas 7 kabupaten/kota, sementara *cluster 2* terdiri atas 31 kabupaten/kota. Secara umum, *cluster 1* memiliki tingkat pengangguran yang lebih tinggi dan rasio pencari kerja terhadap lapangan kerja yang jauh lebih besar dibandingkan *cluster 2*, mencerminkan tantangan dalam penyerapan tenaga kerja. Sebaliknya, *cluster 2* memiliki tingkat partisipasi angkatan kerja yang lebih tinggi dan kondisi pasar tenaga kerja yang lebih stabil, menunjukkan efisiensi dalam menyerap tenaga kerja

**Kata Kunci:** ketenagakerjaan, *hierarchical clustering*, *non-hierarchical clustering*

### 1 PENDAHULUAN

Ketenagakerjaan merupakan salah satu indikator utama dalam menilai kesejahteraan masyarakat dan tingkat pembangunan ekonomi suatu wilayah. Jawa Timur, sebagai salah satu provinsi dengan populasi terbesar di Indonesia, memiliki dinamika ketenagakerjaan yang kompleks. Indikator seperti Tingkat Pengangguran Terbuka (TPT), Tingkat Partisipasi Angkatan Kerja (TPAK), dan distribusi pekerja di berbagai sektor memberikan gambaran yang penting tentang kondisi pasar tenaga kerja di provinsi ini. Menurut (BPS Jatim, 2023) Tingkat Partisipasi Angkatan Kerja (TPAK) Jawa Timur mencapai 72,56% yang menunjukkan bahwa mayoritas penduduk usia kerja aktif dalam kegiatan ekonomi. Dari penduduk angkatan kerja tersebut, terdapat penduduk yang bekerja sebesar 22,70 juta orang, bertambah sekitar 1,09 juta sebesar 5,04% dibandingkan keadaan Agustus 2022.

Di sisi lain, jumlah penduduk usia kerja di Jawa Timur pada Agustus 2023 sebesar 32,90 juta orang. Jumlah tersebut terdiri dari 23,87 juta orang penduduk angkatan kerja dan 9,03 juta orang penduduk bukan angkatan kerja. Jumlah pengangguran terbuka pada Agustus 2023 mencapai sekitar 1,17 juta orang, berkurang 90,13 ribu orang jika dibandingkan keadaan Agustus 2022. Tingkat Pengangguran Terbuka (TPT) Jawa Timur mencapai 4,88%, mengalami penurunan

sebesar 0,61% poin dibandingkan TPT Agustus 2022 sebesar 5,49% (BPS Jatim, 2023). Salah satu karakteristik daerah perkotaan adalah memiliki kecenderungan sebagai tujuan orang untuk menetap atau mencari nafkah di daerah tersebut. Hal ini dikarenakan daerah perkotaan memiliki infrastruktur yang lebih lengkap dan cenderung menjadi pusat aktivitas perekonomian. Karakteristik tersebut menjadi salah satu penyebab TPT perkotaan cenderung lebih tinggi dibandingkan TPT perdesaan. Ketidakseimbangan kondisi ketenagakerjaan ini menunjukkan pentingnya pendekatan analisis yang lebih mendalam. Pendekatan clustering dapat digunakan untuk mengelompokkan wilayah kabupaten/kota di Provinsi Jawa Timur berdasarkan indikator ketenagakerjaan yang serupa. Metode ini memungkinkan pengelompokan daerah dengan tingkat pengangguran, partisipasi kerja, atau struktur ekonomi yang sama, sehingga mempermudah identifikasi wilayah yang memerlukan perhatian khusus. Dengan pendekatan ini, kebijakan ketenagakerjaan dapat dirancang lebih spesifik untuk mengatasi permasalahan di setiap kelompok wilayah.

Data dari (BPS Jatim, 2023) menunjukkan adanya pergeseran struktur ketenagakerjaan di Jawa Timur. Pada tahun 2023, jumlah tenaga kerja di sektor pertanian menurun, sedangkan sektor industri dan jasa mengalami peningkatan. Daerah dengan ekonomi berbasis agraris seringkali menghadapi tantangan dalam menyerap tenaga kerja baru, sementara daerah perkotaan yang mendukung sektor industri dan jasa menunjukkan pertumbuhan yang lebih pesat. Transformasi ini menjadi peluang sekaligus tantangan dalam menciptakan pemerataan lapangan kerja di seluruh wilayah Jawa Timur. Oleh karena itu, penelitian ini bertujuan untuk menganalisis indikator ketenagakerjaan di Jawa Timur tahun 2023 menggunakan metode *clustering*. Analisis yang akan digunakan yaitu metode *hierarchical clustering* dan *non-hierarchical clustering*. Selanjutnya dilakukan pemilihan metode untuk mengetahui metode terbaik yang dapat digunakan dengan menggunakan nilai indeks validitas *cluster*. Hasil pengelompokan ini diharapkan dapat memberikan wawasan yang lebih mendalam dan menyeluruh mengenai karakteristik perkembangan tenaga kerja di Jawa Timur yang memiliki karakteristik serupa sehingga memudahkan perumusan kebijakan ketenagakerjaan lebih efektif, seperti fokus pada sektor-sektor dengan tingkat pengangguran tinggi atau peningkatan keterampilan lulusan pendidikan tertentu. Hasil dari penelitian ini juga diharapkan dapat mendukung upaya pemerataan pembangunan ekonomi dan meningkatkan kesejahteraan tenaga kerja di Jawa Timur.

Beberapa penelitian sebelumnya telah dilakukan analisis mengenai ketenagakerjaan. Sebagai contoh, (Syafiyah dkk, 2022) menggunakan perbandingan *hierarchical* dan *non-hierarchical clustering* pada data indikator ketenagakerjaan di Jawa Barat tahun 2020. Dalam penelitian itu menyajikan perbandingan antara metode *hierarchical clustering*, yaitu *single linkage*, *average linkage*, dan *complete linkage*, dan metode *hierarchical cluster*, yaitu *k-means cluster*. Hasil analisis menunjukkan bahwa metode *complete linkage 2 cluster* adalah metode yang paling cocok dalam pengelompokan kabupaten/kota di Provinsi Jawa Barat berdasarkan indikator ketenagakerjaan. Terdapat dua kelompok yang terbentuk berdasarkan *complete linkage*, yaitu: *cluster 1* yang beranggotakan 18 kabupaten/kota dengan karakteristik perkembangan tenaga kerja yang sudah baik dan *cluster 2* yang beranggotakan 9 kabupaten/kota merupakan *cluster* dengan karakteristik perkembangan tenaga yang belum maksimal. Penelitian lainnya (Novaldi & Wijayanto, 2023) menggunakan perbandingan metode *Agglomerative* dan *K-Mean* untuk menganalisis kualitas pemuda di Indonesia pada tahun 2022. Pada penelitian ini metode terbaik

yang digunakan untuk pengelompokan provinsi di Indonesia berdasarkan data kepemudaan adalah metode *Agglomerative Hierarchical (Ward's method)* dimana jumlah cluster sebanyak 2.

## 2 METODE PENELITIAN

Dalam penelitian ini data yang digunakan adalah data sekunder. Data tersebut diperoleh dari website BPS (Badan Pusat Statistika) dalam Publikasi Keadaan Angkatan Kerja Provinsi Jawa Timur Tahun 2023. Adapun variabel yang diamati dalam penelitian ini adalah sebagai berikut.

**Tabel 1.** Variabel Penelitian

Variabel	Keterangan	Skala
TPT	Tingkat Pengangguran Terbuka (TPT) adalah persentase jumlah pengangguran terhadap jumlah angkatan kerja. TPT yang tinggi menunjukkan bahwa terdapat banyak angkatan kerja yang tidak terserap pada pasar kerja.	Rasio
TPAK	Tingkat Partisipasi Angkatan Kerja (TPAK) adalah persentase banyaknya angkatan kerja terhadap banyaknya penduduk yang berumur sepuluh tahun ke atas. TPAK merupakan indikator yang mengindikasikan besarnya persentase penduduk usia kerja yang aktif secara ekonomi di suatu daerah. Semakin tinggi TPAK, menunjukkan bahwa semakin tinggi pula pasokan tenaga kerja ( <i>labour supply</i> ) yang tersedia untuk memproduksi barang dan jasa suatu perekonomian di kota tersebut.	Rasio
RPLK	Rasio Pencari Kerja terhadap Lapangan Kerja (RPLK) adalah persentase orang yang bekerja terhadap jumlah penduduk yang termasuk angkatan kerja.	Rasio

### 2.1 Analisis Cluster

Analisis *cluster* adalah metode yang tepat untuk mengidentifikasi objek- objek yang homogen ke dalam kelompok-kelompok yang disebut *cluster*. *Cluster-cluster* yang terbentuk memiliki homogenitas internal yang tinggi dan heterogenitas eksternal yang tinggi (Hair et al., 2010). Tujuan pengelompokan adalah agar objek yang bergabung dalam sebuah kelompok merupakan objek-objek yang mirip (atau berhubungan) satu sama lain dan berbeda (tidak berhubungan) dengan objek atau kelompok lain. Lebih besar tingkat kemiripan (atau homogenitas) dalam suatu kelompok dan lebih besar perbedaannya diantara kelompok. Secara garis besar metode *cluster* terbagi menjadi dua, yaitu metode *Hierarchical Clustering* dan *Non-Hierarchical Clustering*.

#### 2.1.1 Hierarchical Clustering

*Hierarchical clustering* adalah metode analisis kelompok yang berusaha untuk membangun sebuah hierarki kelompok. *Hierarchical clustering* dibagi menjadi dua yaitu *Agglomeratif Clustering* dan *Difisive Clustering*. *Agglomeratif Clustering* mengelompokkan data dengan pendekatan bawah atas (*bottom up*). Beberapa metode yang umum digunakan dalam pengelompokan dengan *Agglomeratif Clustering* adalah *single linkage*, *complete linkage* dan *average linkage*.

##### 1. Single Linkage

Metode *single linkage* disebut juga metode pautan tunggal dimana kemiripan dua *cluster* diukur menggunakan jarak terdekat sepasang anggota dari dua *cluster* (Umu Sa'adah, 2021).

Penggunaan jarak terdekat menjadikan *cluster* yang terbentuk cenderung memanjang. Untuk menghitung jarak *cluster* dengan *cluster* lainnya dapat dirumuskan sebagai berikut:

$$d(A, B) = \min\{d(y_i, y_j)\} \quad (1)$$

Dimana nilai  $y_i$  dan  $y_j$  menggambarkan jarak terdekat antara *cluster* A dengan *cluster* B.

## 2. Complete Linkage

Metode *complete linkage* disebut juga metode pautan lengkap dimana kemiripan dua *cluster* ditentukan berdasarkan jarak terjauh dari sepasang anggota kedua *cluster*. Sehingga menyebabkan *cluster* yang terbentuk cenderung padat (Umu Sa'adah, 2021). Untuk menghitung jarak *cluster* dengan *cluster* lainnya dapat dirumuskan sebagai berikut:

$$d(A, B) = \max\{d(y_i, y_j)\} \quad (2)$$

## 3. Average Linkage

Metode *average linkage* disebut juga pautan rata-rata dimana menggunakan rata-rata jarak seluruh objek dalam satu *cluster* dengan jarak seluruh objek pada *cluster* yang lain. Metode ini memberikan hasil yang lebih seimbang dan stabil dibandingkan *single* dan *complete linkage* (Umu Sa'adah, 2021). Untuk menghitung jarak *cluster* dengan *cluster* lainnya dapat dirumuskan sebagai berikut:

$$d(A, B) = \frac{1}{n_A n_B} \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} d(y_i, y_j) \quad (3)$$

Adapun langkah-langkah yang perlu dilakukan untuk melakukan analisis *cluster* dengan metode *hierarchical clustering* adalah sebagai berikut

1. Menghitung jarak antar data menggunakan metode jarak Euclidean untuk mengukur tingkat kemiripan antar objek. Jarak Euclidean dihitung dengan rumus

$$d(A, B) = \sqrt{\sum_{i=1}^n (y_i - y_j)^2} \quad (4)$$

dimana  $y_i$  dan  $y_j$  adalah dua data dengan  $n$  variabel.

2. Memvisualisasikan hasil pengelompokan dalam bentuk dendrogram untuk mempermudah interpretasi berdasarkan jarak Euclidean yang telah dihitung dengan teknik *single linkage*, *complete linkage*, dan *average linkage*.
3. Melakukan evaluasi validitas cluster menggunakan indeks seperti *Connectivity*, *Silhouette*, dan *Dunn*. Pilih jumlah cluster optimal berdasarkan nilai *Connectivity* terendah serta *Silhouette* dan *Dunn* tertinggi.
4. Menentukan metode *hierarchical clustering* terbaik berdasarkan koefisien aglomerasi dengan memilih nilai koefisien aglomerasi yang mendekati 1 sebagai metode yang optimal untuk pengelompokan.
5. Melakukan interpretasi hasil analisis *clustering*.

### 2.1.2 Non-Hierarchical Clustering

*Non-Hierarchical clustering* adalah salah satu metode pengelompokan dimana objek akan dikelompokkan ke dalam kelompok yang telah ditentukan sebelumnya. Metode *non-hierarchical clustering* yang paling umum digunakan adalah *k-means*.

Pengelompokkan data menggunakan *k-means* dilakukan dengan mempartisi data yang ada kedalam kelompok sehingga data berkarakteristik sama dimasukkan kedalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok lain (Prasetyo, 2014). Menurut (MacQueen, 1967), metode ini dimulai dengan menentukan jumlah *cluster* terlebih dahulu yang diinginkan misalkan *c cluster*.

Adapun langkah-langkah yang perlu dilakukan untuk melakukan analisis *cluster* dengan metode *non-hierarchical clustering* dengan teknik *k-means* adalah sebagai berikut

1. Menentukan banyaknya kluster (*k*) menggunakan metode elbow. Metode ini memakai plot antara *number of cluster* dan *total within sum of square*, dimana jumlah *cluster* optimum yang terbentuk ditandai dengan grafik yang mulai melandai.
2. Menggunakan *silhouette plot* untuk memvalidasi jumlah cluster yang dipilih. Jumlah *cluster* optimal dipilih berdasarkan rata-rata nilai *silhouette* tertinggi atau titik plot yang paling tinggi atau menukik.
3. Melakukan inisialisasi centroid awal secara acak untuk masing-masing cluster, kemudian hitung jarak Euclidean antara setiap data dan centroid dengan rumus:

$$d(A, c) = \sqrt{\sum_{i=1}^n (y_i - c)^2} \quad (5)$$

dimana *c* adalah centroid cluster

4. Mengelompokkan setiap data ke cluster dengan jarak terdekat ke centroid, kemudian hitung ulang centroid dengan mengambil rata-rata posisi dari semua data dalam cluster tersebut.
5. Mengevaluasi hasil clustering menggunakan indeks validasi seperti *Connectivity*, *Silhouette*, dan *Dunn*. Pilih hasil clustering dengan nilai indeks validasi terbaik untuk memastikan kualitas *cluster*

### 3 HASIL DAN PEMBAHASAN

#### 3.1 Karakteristik Data

Karakteristik data digunakan untuk menunjukkan gambaran umum tiap variabel yang akan diuji dalam menganalisis indikator ketenagakerjaan di Jawa Timur pada tahun 2023.

**Tabel 2.** Statistika Deskriptif

Variabel	N	Minimum	Median	Maximum	Mean
TPT	38	1.71	4.6650	8.05	4.6629
TPAK	38	66.89	72.4950	81.64	73.1592
RPLK	38	4.04	22.2650	176.61	40.4133

Tabel 2 menunjukkan bahwa dari 38 kabupaten/kota di Jawa Timur rata-rata TPT (Tingkat Pengangguran Terbuka) memiliki distribusi yang stabil dengan variasi yang kecil, mencerminkan tingkat pengangguran yang relatif seragam. TPAK (Tingkat Partisipasi Angkatan Kerja) menunjukkan tingkat partisipasi yang tinggi dan konsisten, dengan sebaran data yang merata. Sebaliknya, RPLK (Rasio Penduduk Luar Kota) memiliki variasi yang sangat besar, menunjukkan adanya perbedaan yang signifikan di antara data, dengan beberapa nilai jauh lebih tinggi dibandingkan yang lain. Hal ini menandakan bahwa stabilitas distribusi lebih terlihat pada TPT dan TPAK dibandingkan RPLK.

### 3.2 Pemeriksaan Uji Asumsi

Pendeteksi asumsi berupa uji homogenitas yaitu untuk mengetahui apakah variasi beberapa data dari populasi memiliki varians yang sama atau tidak. Metode yang digunakan yaitu *Bartlett's Test*. Selain itu pengujian KMO juga perlu dilakukan untuk melihat apakah data telah memenuhi asumsi kecukupan sampel sehingga bisa dianalisis lebih lanjut.

**Tabel 3.** KMO and Bartlett's Test

Uji		Nilai uji
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.618
Bartlett's Test of Sphericity	Approx. Chi-Square Df	18.243 3
	Sig.	.000

Tabel 3 menunjukkan nilai KMO sebesar 0.618 yang lebih besar dari 0.5, sehingga dapat diimpulkan bahwa sampel yang digunakan dalam penelitian telah cukup untuk menggambarkan populasi yang ada di Jawa Timur. Pada Uji *Bartlett* menunjukkan hasil yang signifikan dimana nilai Sig.  $0.000 < 0.05$ , sehingga matriks korelasi berbeda signifikan dari matriks identitas. Hal ini menunjukkan bahwa variabel memiliki hubungan yang cukup untuk dianalisis lebih lanjut dengan analisis faktor

**Tabel 4.** Loading Factor

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
PC1	1.831	61.022	61.022
PC2	.714	23.794	<b>84.816</b>
PC3	.456	15.184	100.000

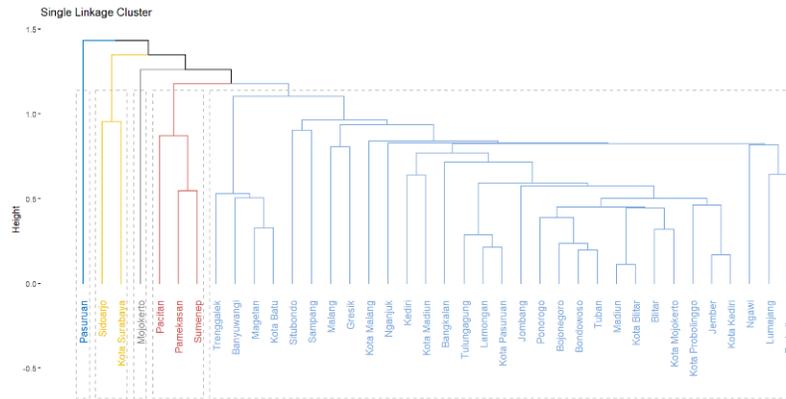
Tabel 4 menunjukkan bahwa komponen utama pertama (PC1) memiliki nilai eigenvalue sebesar 1.831, yang menjelaskan sekitar 61.022% varians total dalam data. Komponen kedua (PC2) memiliki nilai eigenvalue 0.714 yang menjelaskan tambahan 23.794% varians, sehingga kedua komponen secara kumulatif menjelaskan 84.816% varians. Berdasarkan Kriteria Kaiser dimana jumlah *principal components* kumulatif berada diantara 70 hingga 90% maka PC1 dan PC2 dianggap signifikan. Oleh karena itu, dua komponen utama ini dianggap cukup untuk merangkum informasi dari data awal, dengan total varians kumulatif yang dijelaskan mencapai 84.816%, yang sudah cukup memadai untuk interpretasi lebih lanjut.

### 3.3 Metode Clustering

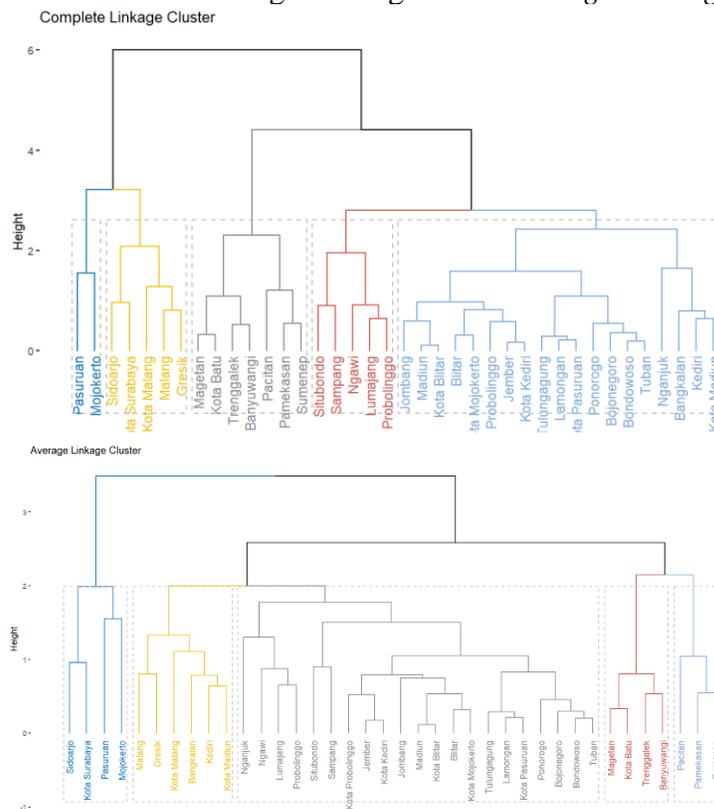
Metode *clustering* yang digunakan dalam penelitian ini meliputi metode *hierarchical cluster* dan *non-hierarchical cluster*

#### 3.1.1 Metode Hierarchical Cluster

*Hierarchical clustering* adalah metode analisis data yang bertujuan mengelompokkan objek ke dalam cluster berdasarkan tingkat kemiripan atau jarak antar data. Metode ini membangun hierarki pengelompokan yang divisualisasikan dalam bentuk dendrogram. *Hierarchical clustering* tidak memerlukan penentuan jumlah cluster (k) di awal, sehingga cocok untuk eksplorasi data dengan struktur yang kompleks.



**Gambar 1.** Dendrogram dengan metode *Single Linkage*



**Gambar 2.** Dendrogram dengan metode *Complete Linkage* dan *Average Complete*

Gambar 1 dan 2 menjelaskan bahwa dengan metode umum *Linkage* yang sering digunakan yaitu meliputi *single linkage*, *complete linkage* dan *average linkage*, maka data indikator ketenagakerjaan di Jawa Timur Tahun 2023 dapat dikelompokkan menjadi 2, 3, 4, dan 5 cluster. Selain itu, evaluasi dengan matrik validitas *cluster*, seperti *Connectivity*, *Silhouette* atau *Dunn*, dapat digunakan untuk menilai kualitas cluster yang dihasilkan oleh setiap metode linkage.

**Tabel 5.** Indeks Validasi *Hierarchical Cluster*

Metode	Cluster	Connectivity	Silhoutte	Dunn
Single Linkage	2	<b>3.9623</b>	0.3930	0.2386
	3	8.0369	<b>0.4088</b>	0.2612
	4	10.5659	0.2497	0.2439
	5	15.2972	0.2735	<b>0.3009</b>
Complete Linkage	2	<b>7.8861</b>	<b>0.4317</b>	0.1910
	3	16.1921	0.3828	0.2620
	4	20.7683	0.3536	0.3003
	5	28.8798	0.2992	<b>0.3403</b>
Average Linkage	2	<b>7.1325</b>	<b>0.4413</b>	0.2439
	3	15.4385	0.3555	<b>0.3113</b>
	4	16.8159	0.3367	0.3113
	5	24.3841	0.3402	0.2552

Tabel 5 menunjukkan bahwa metode *hierarchical clustering* dengan menggunakan teknik *single linkage*, *complete linkage* dan *average linkage* menghasilkan jumlah *cluster* yang sama yaitu sebanyak 2 *cluster* sebagai metode pengelompokan yang terbaik. Pemilihan jumlah *cluster* dilihat dari nilai *Connectivity* terendah serta nilai *Silhoutte* dan *Dunn* tertinggi. Selanjutnya untuk menentukan teknik mana yang terbaik perlu dilihat nilai koefisien aglomerasi. Nilai dengan koefisien aglomerasi mendekati 1 dapat dipilih sebagai teknik untuk melakukan *clustering*.

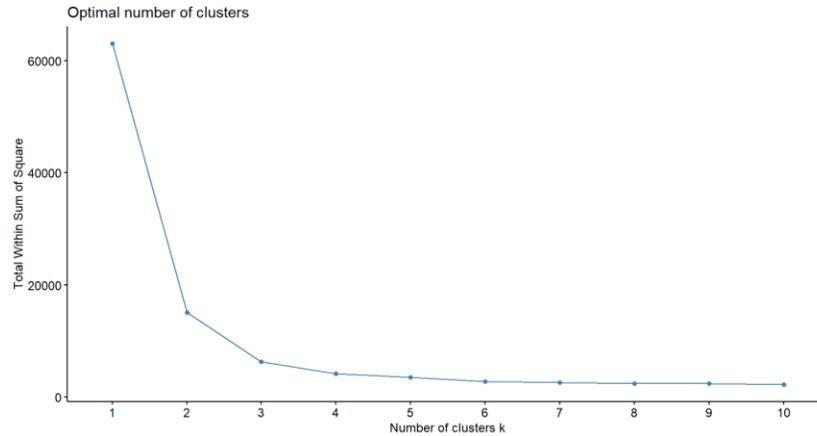
**Tabel 6.** Koefisien Aglomerasi

Metode	Koefisien Aglomerasi
<i>Single Linkage</i>	0.6665
<i>Complete Linkage</i>	<b>0.9190</b>
<i>Average Linkage</i>	0.9170

Tabel 6 menunjukkan bahwa nilai koefisien aglomerasi yang paling besar atau mendekati 1 adalah metode *complete linkage* dengan nilai koefisien aglomerasi sebesar 0.9190, sehingga *complete linkage* dapat dipilih sebagai metode dengan hasil *clustering* terbaik.

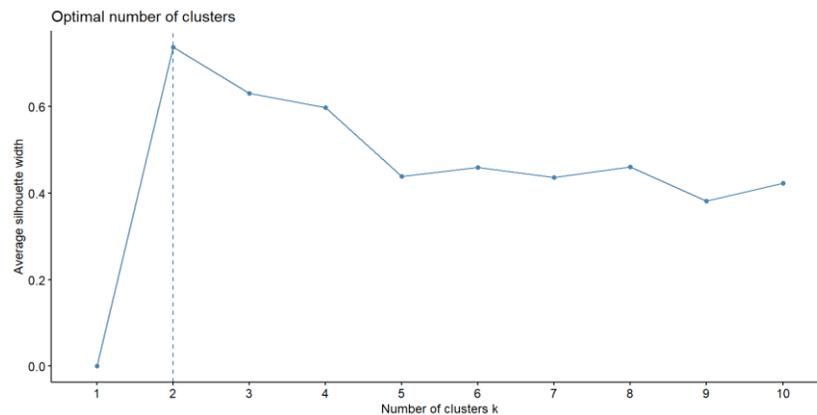
### 3.1.2 Metode *Non-Hierarchical Cluster*

*Hierarchical clustering* adalah metode analisis data yang bertujuan mengelompokkan objek ke dalam cluster berdasarkan tingkat kemiripan atau jarak antar data. Metode ini membangun hierarki



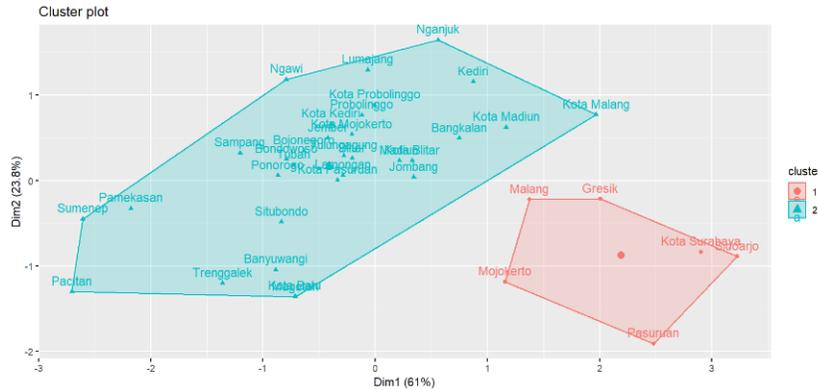
**Gambar 2.** Penentuan Jumlah *Cluster* menggunakan Metode *Elbow Plot*

Gambar 2 pada visualisasi jumlah cluster menggunakan *elbow plot* menunjukkan penurunan signifikan hingga  $k = 2$  dan setelah itu grafik melandai, maka jumlah *cluster* optimal adalah 2. Menambah *cluster* lebih dari 2 tidak efisien karena perbaikan kualitas *clusterisasi* minimal dibandingkan dengan kenaikan jumlah *cluster*.



**Gambar 3.** Penentuan Jumlah *Cluster* menggunakan Metode *Silhouette Plot*

Gambar 3 menunjukkan nilai *silhouette width* tertinggi berada pada saat jumlah *cluster* sebanyak 2 *cluster* dengan sebagian besar titik data memiliki nilai di atas 0.4 dan hanya sedikit yang mendekati 0. Sehingga  $k = 2$  dapat dianggap sebagai jumlah *cluster* yang memadai. Oleh karena itu, berdasarkan hasil yang diperoleh dari metode *elbow plot* dan *silhouette plot*, maka dapat disimpulkan bahwa jumlah *cluster* optimum dengan menggunakan metode *k-means cluster* adalah sebanyak 2 *cluster*. Hasil pengelompokkan menggunakan *k-means cluster* adalah sebagai berikut.



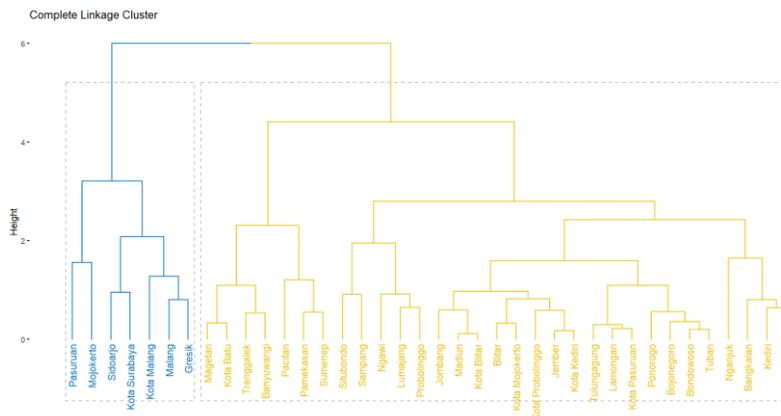
**Gambar 4.** Visualisasi Pengelompokan menggunakan *K-Means Cluster* dengan  $k=2$

Berdasarkan Gambar 4 diketahui bahwa *cluster 1* terdiri dari 6 kabupaten/kota dan *cluster 2* terdiri dari 32 kab/kota di Provinsi Jawa Timur.

**Tabel 7.** Perbandingan Indeks Validitas *Hierarchical Cluster* Terbaik dan *Non-Hierarchical Cluster*

Metode	Connectivity	Silhouette	Dunn
Hierarchical <i>Complete Linkage</i>	<b>7.8861</b>	<b>0.4317</b>	<b>0.1910</b>
<i>K-Means</i>	9.1333	0.4154	0.1454

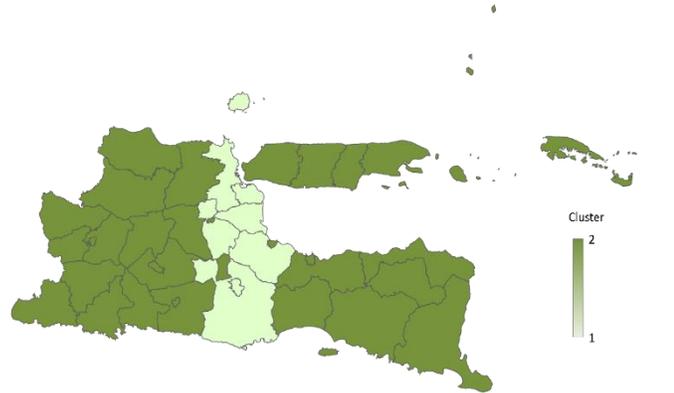
Tabel 7 menunjukkan bahwa metode *hierarchical complete linkage* memiliki nilai validitas metode *connectivity* terkecil serta *silhouette* dan *Dunn* terbesar. Oleh karena itu, metode terbaik dalam penelitian ini dengan menggunakan data indeks ketenagakerjaan di Jawa Timur tahun 2023 adalah metode *hierarchical complete linkage* dengan jumlah *cluster* sebanyak 2 *cluster*.



**Gambar 5.** Dendrogram metode *Complete Linkage 2 Cluster*

Gambar 5 menunjukkan dendrogram metode *Complete Linkage* yang mengelompokan wilayah berdasarkan kemiripan karakteristiknya. Pada ketinggian sekitar 6, wilayah terbagi menjadi dua *cluster* utama. *Cluster* pertama (berwarna biru) mencakup wilayah Pasuruan, Mojokerto, Sidoarjo, Kota Surabaya, Kota Malang, Malang, dan Gresik, yang memiliki kemiripan lebih besar di antara mereka. *Cluster* kedua (berwarna kuning) terdiri dari wilayah yang lebih beragam, seperti

Banyuwangi, Jember, Probolinggo, Lamongan, dan lain-lain, yang menunjukkan variasi karakteristik yang lebih signifikan. *Subcluster* di dalam masing-masing *cluster* menggambarkan kedekatan karakteristik antarwilayah tertentu, seperti Pasuruan dan Mojokerto di *cluster* biru, atau Probolinggo dan Jombang di *cluster* kuning. Pengelompokan ini mencerminkan hubungan hierarkis antarwilayah berdasarkan karakteristik tertentu, dengan *cluster* yang berbeda menunjukkan tingkat variasi yang signifikan.



**Gambar 6.** Peta Penyebaran *Cluster* Indikator Ketenagakerjaan di Jawa Timur

Gambar 6 divisualisasikan dengan dua warna yang berbeda. Peta penyebaran *cluster* di Jawa Timur menunjukkan dua kelompok berdasarkan indikator ketenagakerjaan. *Cluster* 1 (hijau muda) mencakup wilayah pusat seperti Kota Surabaya dan Sidoarjo, yang cenderung lebih urban dengan akses ketenagakerjaan yang lebih baik. *Cluster* 2 (hijau tua) meliputi sebagian besar wilayah lainnya, yang didominasi oleh sektor agraris dan keterbatasan akses kerja. Pola ini mencerminkan perbedaan ketenagakerjaan antara wilayah urban dan rural di Jawa Timur.

**Tabel 8.** Hasil Pembentukan *Cluster*

<i>Cluster</i>	Anggota
1	Pasuruan, Mojokerto, Sidoarjo, Kota Surabaya, Kota Malang, Malang, Gresik, Magetan, Kota Batu, Trenggalek, Banyuwangi, Pacitan, Pamekasan, Sumenep, Situbondo, Sampang, Ngawi, Lumajang, Probolinggo, Jombang, Madiun, Kota Blitar, Kota Mojokerto, Kota Probolinggo, Jember, Kota Kediri, Tulungagung, Lamongan, Kota Pasuruan, Ponorogo, Bojonegoro, Bondowoso, Tuban, Nganjuk, Bangkalan, Kediri, Kota Madiun
2	

Pembentukan *cluster* pada Tabel 8, mencerminkan adanya perbedaan karakteristik sosial-ekonomi dan geografis di wilayah Jawa Timur. *Clusterisasi* ini dapat digunakan sebagai dasar untuk merancang kebijakan pembangunan yang lebih terarah, misalnya untuk meningkatkan infrastruktur atau peluang ekonomi di wilayah *cluster* 2.

**Tabel 9.** Rata-rata variabel

<i>Cluster</i>	TPT	TPAK	RPLK
1	5.354286	70.06143	112.9727
2	4.287419	73.85871	24.02634

Hasil *cluster* dari indikator ketenagakerjaan di Jawa Timur tahun 2023 sebagai berikut:

- a. *Cluster 1*, rata-rata TPT sebesar 5,35% menunjukkan tingkat pengangguran yang relatif lebih tinggi dibandingkan *Cluster 2*, meskipun TPAK sebesar 70,06% menunjukkan partisipasi angkatan kerja yang cukup baik. Namun, nilai RPLK sebesar 112,97 menunjukkan bahwa jumlah pencari kerja jauh melebihi jumlah lapangan kerja yang tersedia, yang bisa menjadi indikasi tingginya tekanan pasar kerja
- b. *Cluster 2* memiliki rata-rata TPT lebih rendah, yaitu 4,28%, yang mencerminkan tingkat pengangguran yang lebih kecil. TPAK sebesar 73,86% menunjukkan partisipasi angkatan kerja yang lebih tinggi dibanding *Cluster 1*. Namun, nilai RPLK yang jauh lebih rendah, yaitu 24,03, menunjukkan bahwa jumlah pencari kerja relatif lebih kecil dibandingkan dengan lapangan kerja yang tersedia, yang dapat mencerminkan keseimbangan pasar kerja yang lebih baik.

Berdasarkan analisis ketenagakerjaan, diperlukan kebijakan yang berbeda untuk meningkatkan kondisi di setiap *cluster*. Pada *cluster 1*, yang memiliki Tingkat Pengangguran Terbuka (TPT) tinggi dan Rasio Pencari Kerja terhadap Lapangan Kerja (RPLK) yang besar, diperlukan upaya untuk menyediakan lebih banyak pelatihan keterampilan yang sesuai dengan kebutuhan industri lokal serta mendorong investasi guna menciptakan lapangan kerja baru. Selain itu, sinergi antara pemerintah, lembaga pendidikan, dan perusahaan perlu ditingkatkan untuk memastikan kesiapan lulusan memasuki pasar kerja, serta menyediakan informasi lowongan pekerjaan yang lebih mudah diakses melalui bursa kerja atau platform online.

Di sisi lain, *cluster 2*, dengan TPT rendah dan Tingkat Partisipasi Angkatan Kerja (TPAK) tinggi, memerlukan kebijakan untuk meningkatkan keterlibatan kelompok tertentu, seperti perempuan dan kaum muda, melalui pelatihan kerja atau dukungan wirausaha. Selain itu, pemanfaatan kondisi pasar kerja yang stabil dapat dilakukan dengan mendorong inovasi dan ekspansi sektor ekonomi utama untuk meningkatkan pertumbuhan ekonomi lokal. Secara umum, pembangunan infrastruktur yang menunjang kegiatan ekonomi, peningkatan kualitas pendidikan, dan pengembangan ekonomi kreatif berbasis digital juga menjadi kebijakan strategis yang diperlukan di kedua *cluster*. Dengan kebijakan ini, diharapkan ketimpangan ketenagakerjaan dapat diminimalkan, serta mendukung pertumbuhan ekonomi yang inklusif dan berkelanjutan.

#### 4 KESIMPULAN

Pengujian *hierarchical clustering* dan *non-hierarchical clustering* menggunakan data indikator ketenagakerjaan di Jawa Timur tahun 2023 ditemukan bahwa metode *clustering* paling baik adalah metode *hierarchical clustering (complete linkage)*. Pemilihan metode *clustering* terbaik yaitu dengan melihat nilai validitas metode *connectivity* terkecil serta *silhouette* dan *dunn* terbesar. Diperoleh jumlah *cluster* sebanyak 2 *cluster* dimana *cluster 1* terdiri atas 7 kabupaten/kota, sementara *cluster 2* terdiri atas 31 kabupaten/kota. Secara umum, kabupaten/kota menunjukkan perbedaan signifikan dalam kondisi ketenagakerjaan antara dua *cluster*. *Cluster 1* memiliki tingkat pengangguran yang lebih tinggi dan rasio pencari kerja terhadap lapangan kerja yang jauh lebih besar dibandingkan *Cluster 2*, mencerminkan tantangan dalam penyerapan tenaga kerja. Sebaliknya, *Cluster 2* memiliki tingkat partisipasi angkatan kerja yang lebih tinggi dan kondisi pasar tenaga kerja yang lebih stabil, menunjukkan efisiensi dalam menyerap tenaga kerja. Perbedaan ini menggambarkan variasi dinamika pasar kerja di setiap *cluster*. Diperlukan kebijakan yang berbeda untuk meningkatkan kondisi di setiap *cluster*. Untuk *Cluster 1*, kebijakan harus

difokuskan pada peningkatan pelatihan kerja, penciptaan lapangan kerja baru melalui investasi, dan sinergi antara pemerintah, pendidikan, serta industri. Sementara itu, di *Cluster 2*, kebijakan harus mendukung keterlibatan kelompok tertentu, seperti perempuan dan pemuda, melalui pelatihan kerja dan program kewirausahaan, serta mendorong inovasi dan ekspansi sektor ekonomi utama.

### UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah mendukung terlaksananya penelitian ini.

### DAFTAR PUSTAKA

- BPS Jatim. (2023). *Keadaan Angkatan Kerja Provinsi Jawa Timur*. <https://jatim.bps.go.id/id>
- MacQueen, J. . (1967). Some Methods For Classification and Analysis of Multivariate Observations. *5-Th Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, University of California Press. (Pp. 1:281-297).
- Novaldi, J., & Wijayanto, A. W. (2023). Analisis Cluster Kualitas Pemuda di Indonesia pada Tahun 2022 dengan Agglomerative Hierarchical dan K-Means. *Komputika : Jurnal Sistem Komputer*, 12(2), 91–99. <https://doi.org/10.34010/komputika.v12i2.10348>
- Prasetyo, E. (2014). *Data mining Mengolah Data menjadi Informasi Menggunakan Matlab*. CV. Andi Offset.
- Syafiyah, U., Puspitasari, D. P., Asrafi, I., Wicaksono, B., & Sirait, F. M. (2022). Analisis Perbandingan Hierarchical dan Non-Hierarchical Clustering Pada Data Indikator Ketenagakerjaan di Jawa Barat Tahun 2020. *Seminar Nasional Official Statistics*, 2022(1), 803–812. <https://doi.org/10.34123/semnasoffstat.v2022i1.1221>