

Pengelompokan Kabupaten dan Kota di Jawa Timur Berdasarkan Percepatan Pemulihan Ekonomi Menggunakan Pendekatan Hierearki

Clustering of Regencies and Cities in East Java based on the Acceleration of Economic Recovery Using a Hierarchical Approach

Arinda Mahadesyawardani¹, Azizah Atsariyyah Zhafira², Jovansha Ariyawan³, Edla Putri Humaira⁴, M. Fariz Fadillah Mardianto^{5a)}, Dita Amelia⁶, Elly Ana⁷

^{1, 2, 3, 4, 5, 6, 7} Program Studi S1 Statistika, Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga Surabaya

^{a)} Corresponding author: m.fariz.fadillah.m@fst.unair.ac.id

ABSTRACT

The Covid-19 pandemic's diverse impact on Indonesia's economy, particularly in East Java, spurred the government to formulate a comprehensive work plan targeting three key objectives, one of which is to expedite economic recovery. This plan focuses on three key indicators: economic growth, open unemployment rate (TPT), and the gini ratio. It is known that during the pandemic, East Java initially experienced economic growth that contracted until eventually showing positive growth in the second quarter of 2021, which has been supported by national policies. This study explores district and city classification in East Java based on economic recovery indicators through hierarchical clustering. The analysis identifies Ward's linkage as the most effective model, with a cophenetic correlation coefficient of 0.9311. Internal clustering validation tests reveal two optimal clusters. Cluster 1 is characterized by a notably high average acceleration of economic recovery across all three indicators. The findings suggest that the government should optimize the economic stimulus program for cluster 2 and focus on enhancing income redistribution and job opportunities for cluster 1.

Keywords: *hierarchical clustering, internal clustering validation, cophenetic correlation, pemulihan ekonomi nasional*

1. Pendahuluan

Krisis yang terjadi pada sektor ekonomi, sosial, dan kesehatan karena pandemi Covid-19 menimbulkan adanya pembatasan kegiatan masyarakat di berbagai wilayah sebagai langkah pencegahan penyebaran pandemi kala itu sehingga terjadi penurunan yang signifikan salah satunya dalam sektor ekonomi, baik secara global maupun nasional, termasuk Jawa Timur. Pertumbuhan ekonomi yang menurun secara drastis menciptakan tantangan yang signifikan bagi pemulihan ekonomi. Hal ini mendorong pemerintah untuk menyeimbangkan serta mengoptimalkan kebijakan untuk memulihkan ekonomi dengan menetapkan rencana kerja mengenai sasaran utama pembangunan.

Dalam upaya membangkitkan keadaan ekonomi pasca pandemi Covid-19, pemerintah menetapkan rencana kerja dengan tiga sasaran utama yang salah satunya adalah percepatan pemulihan ekonomi. Dalam sasaran utama pembangunan tersebut telah ditetapkan indikator yang diantaranya tingkat pertumbuhan ekonomi, tingkat pengangguran terbuka, dan rasio gini sebagai fokus utama atau perhatian khusus yang dapat menggambarkan keadaan ekonomi Indonesia. Pemerintah menetapkan percepatan pemulihan ekonomi sebagai salah satu prioritas utama setelah mengalami guncangan akibat pandemi Covid-19. Pemerintah daerah juga memiliki peran penting atau krusial dalam mendukung upaya percepatan pemulihan ekonomi tersebut. Langkah-langkah strategis dan kebijakan ekonomi yang diterapkan menjadi kunci dalam merespons ketidakpastian ini.

Nilai pertumbuhan ekonomi Jawa Timur pada tahun 2020 mengalami penurunan hingga mencapai angka -2,33% (Badan Pusat Statistik Jawa Timur, 2022). Hal ini menyebabkan Jawa Timur juga turut serta dalam upaya percepatan pemulihan ekonomi. Pemulihan ekonomi di Jawa Timur didukung oleh kebijakan pemulihan ekonomi nasional (PEN) yang dilaksanakan oleh pemerintah pusat. Namun, upaya pemulihan ekonomi ini masih mengalami kendala karena kabupaten atau kota di Jawa Timur masih memiliki ketimpangan atau terdapat kesenjangan yang tidak homogen dalam pengambilan kebijakan. Dalam mengatasi kendala tersebut, penting bagi pemerintah Provinsi Jawa Timur untuk melakukan pengelompokan atau *clustering* wilayah berdasarkan indikator percepatan pemulihan ekonomi sebelum pengambilan kebijakan. Dengan metode *clustering* ini, peneliti dapat mengidentifikasi pola-pola atau karakteristik bersama dalam kelompok-kelompok ekonomi di Jawa Timur sehingga diharapkan pemerintah dapat merancang kebijakan yang lebih spesifik dan efektif dalam mendukung percepatan pemulihan ekonomi di tingkat lokal. Melalui penelitian ini, diharapkan dapat ditemukan solusi yang inovatif dan berkelanjutan untuk mendukung pertumbuhan ekonomi yang inklusif di Jawa Timur.

2. Tinjauan Pustaka

2.1. Analisis Cluster

Analisis kelompok (*cluster*) ialah pendekatan untuk mengelompokkan data berdasarkan kesamaan atau hubungan antara objek, menggunakan informasi yang terdapat dalam data untuk menghubungkan objek yang serupa dalam satu kelompok dan membedakannya dari objek dalam kelompok lain (Prasetyo, 2012). Terdapat dua prosedur peng-*clusteran*, salah satunya adalah metode hirarki, yang mengelompokkan objek berdasarkan tingkat kemiripan sehingga objek yang mirip dikelompokkan bersama dalam satu kelompok (Hair dkk., 2010).

2.2. Asumsi Analisis Cluster

Analisis *cluster* mempunyai dua asumsi yang perlu terpenuhi, yakni sampel yang representatif dan adanya multikolinearitas antara setiap variabel (Hair dkk., 2014).

a. Sampel Representatif

Merupakan sampel yang secara kuat mencerminkan atau mewakili karakteristik dari keseluruhan populasi. Untuk menguji representatif sampel dapat dilakukan dengan menggunakan uji KMO, Uji KMO bertujuan untuk mengukur kelayakan sampel secara keseluruhan dan memberikan penilaian terhadap kelayakan sampel untuk setiap indikator secara terpisah. Sampel dianggap mewakili populasi atau representatif ketika nilai Koefisien Keanggotaan KMO berada dalam rentang 0,5 hingga 1. Dirumuskan sebagai berikut (Widarjono, 2010) :

$$KMO = \frac{\sum_{j=1}^p \sum_{k=1}^p r_{x_j x_k}^2}{\sum_{j=1}^p \sum_{k=1}^p r_{x_j x_k}^2 + \sum_{j=1}^p \sum_{k=1}^p r_{x_j x_k, x_l}^2}, \text{ dimana } r_{x_j x_k} = \frac{\frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n}} \sqrt{\frac{\sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}{n}}}$$

Dengan,

$r_{x_j x_k}$: korelasi variabel X_j dan X_k

\bar{x}_j : rerata variabel X_j

\bar{x}_k : rerata variabel X_k

n : banyak pengamatan

b. Multikolinearitas

Adanya multikolinearitas menandakan korelasi linear yang kuat pada beberapa atau semua variabel yang dapat diidentifikasi menggunakan koefisien korelasi antar variabel prediktor. Jika nilai koefisien korelasi melebihi 0,8 maka variabel prediktor pada data amatan tersebut menunjukkan indikasi adanya multikolinearitas dalam model regresi (Naes dkk., 2002).

2.3. Pengelompokan Cluster Hierarki

Ada dua jenis dalam metode hierarki, yakni *agglomerative* (pemusatan) dan *divisive* (pemisahan) dengan mengukur *Euclidean distance* (jarak Euclid) dari objek ke- i menuju objek ke- j didefinisikan sebagai berikut (Afifi dkk., 2020; Wierzchon & Klopotek, 2018).

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2}$$

Hasil dari pembentukan *cluster* melalui pendekatan hierarki dapat ditampilkan melalui *tree diagram* yang sering disebut sebagai dendrogram. Dalam penelitian ini menggunakan teknik pengelompokan *agglomerative* yang diterapkan pada model *single*, *complete*, *average*, dan *ward's linkage* (Wierzchon & Klopotek, 2018).

1. Single Linkage

Model ini mengukur jarak antara dua *cluster* dengan menggunakan jarak minimum atau tautan terpendek antara satu objek dalam *cluster* pertama dengan satu objek dalam *cluster* kedua yang dihitung menggunakan:

$$d_{(uv)w} = \min\{d_{uw}, d_{vw}\}$$

Dengan,

d_{uw} : jarak paling dekat dari *cluster* U dan W

d_{vw} : jarak paling dekat dari *cluster* V dan W

2. Complete Linkage

Model ini mengukur jarak antara dua *cluster* yang menggunakan jarak maksimum atau tautan terjauh antara satu objek dalam *cluster* pertama dengan satu objek dalam *cluster* kedua yang dihitung menggunakan:

$$d_{(uv)w} = \max\{d_{uw}, d_{vw}\}$$

Dengan,

d_{uw} : jarak paling jauh dari *cluster* U dan W

d_{vw} : jarak paling jauh dari *cluster* V dan W

3. *Average Linkage*

Model ini mengukur jarak antara dua *cluster* dengan menggunakan rata-rata dari semua jarak antara objek-objek dalam *cluster* pertama dan *cluster* kedua yang dihitung menggunakan:

$$d_{(uv)w} = \text{avg}\{d_{uw}, d_{vw}\} = \frac{\sum_{i=1}^n \sum_{j=1}^n d_{ij}}{N_{(uv)}N_w}$$

Dengan,

d_{ij} : jarak antara objek ke- i pada *cluster* UV dan objek ke- j pada *cluster* W

$N_{(uv)}N_w$: jumlah objek pada *cluster* UV dan W

4. *Ward's*

Model ini mengukur jarak diantara dua *cluster* dengan meminimalkan peningkatan jumlah kuadrat dalam varian di dalam *cluster* yang baru terbentuk setelah penggabungan, yang dihitung menggunakan:

$$ESS = \sum_{i=1}^n (x_i - \bar{x})^T (x_i - \bar{x})$$

Dengan,

ESS : *Error Sum of Square*

x_i : Vektor objek ke- i

\bar{x} : Vektor rerata semua objek

n : banyak objek

2.4. **Pemilihan Model dan Jumlah Cluster Terbaik**

Cluster yang optimal memiliki tingkat homogenitas yang tinggi serta memiliki perbedaan antara satu *cluster* dengan *cluster* lainnya yang signifikan (Afifi dkk., 2020; Wierzchon & Klopotek, 2018). Beberapa teknik untuk mengukur tingkat kebaikan dari *cluster* adalah koefisien *Cophenetic* dan *Internal Cluster Validation*.

a. *Koefisien Cophenetic*

Koefisien *Cophenetic* merupakan matriks korelasi yang menilai hubungan antara elemen dalam matriks ketidakmiripan asli dan elemen yang berasal dari dendrogram (Silva & Dias, 2013). Formula matematikanya menurut Saracli dkk. (2013), adalah sebagai berikut:

$$r_{coph} = \frac{\sum_{i < j} (d_{ij} - \bar{d}) (d_{cij} - \bar{d}_c)}{\sqrt{[\sum_{i < j} (d_{ij} - \bar{d})^2] [\sum_{i < j} (d_{cij} - \bar{d}_c)^2]}}$$

Dengan,

r_{coph} : koefisien korelasi *cophenetic*

d_{ij} : jarak *euclid* antara objek i dan j

\bar{d} : rerata jarak *euclid* antara objek i dan j

d_{cij} : jarak *cophenetic* antara objek i dan j

\bar{d}_c : rerata jarak *cophenetic* antara objek i dan j

Rentang nilai koefisien *cophenetic* berada diantara -1 hingga 1. Semakin mendekati nilai 1, maka *clustering* dianggap baik.

b. *Internal Cluster Validation*

Validasi internal melibatkan informasi internal dalam data untuk mengevaluasi kualitas *clustering*, mencerminkan kepadatan, mencerminkan hubungan, serta indikasi pemisahan antar *cluster*. Kepadatan terkait dengan homogenitas *cluster* yang dievaluasi melalui varian *intra-cluster*, hubungan mencerminkan penempatan data dalam *cluster* dengan mengukur konektivitas, sementara pemisahan antar *cluster* mengindikasikan jarak antara dua *cluster* (Nahdliyah dkk, 2019).

1. *Indeks Dunn*

Menurut Bezdek, (1995), *Dunn Index* merupakan fungsi validitas yang menilai keefektifan dalam penerapan metode *clustering* yang berbeda, rumus *Dunn Index* adalah sebagai berikut :

$$D = \min_{1 \leq i \leq n} \left\{ \min_{1 \leq k \leq n, i \neq j} \left\{ \frac{d(c_i, c_j)}{\max_{1 \leq k \leq n} (d'(c_k))} \right\} \right\}$$

Dengan,

D : indeks *Dunn*

$d(c_i, c_j)$: jarak antar kelompok c_i dan c_j

$d'(c_k)$: jarak dalam kelompok c_k

Rentang *Dunn Index* adalah 0 sampai 1. Kualitas *cluster* meningkat seiring dengan peningkatan nilai *Dunn Index*.

II. Indeks *Connectivity*

Indeks *connectivity* adalah indeks yang menyatakan tingkat hubungan antar *cluster* yang ditentukan dengan jumlah unit observasi paling dekat. Pembentukan *cluster* yang baik atau optimal ditunjukkan dengan nilai indeks *connectivity* yang semakin kecil (Halim & Widodo, 2017), dengan perhitungan rumus sebagai berikut:

$$Conn = \sum_{i=1}^N \sum_{j=1}^L x_{i,nn_{i(j)}}$$

Dengan,

$Conn$: Indeks *Connectivity*

$nn_{i(j)}$: pengamatan tetangga paling dekat dari data ke- j ke data ke- i

N : banyak pengamatan

L : banyak *cluster*

III. Indeks *Silhouette*

Indeks *Silhouette* mengukur posisi tiap objek dalam *cluster* dengan menilai perbandingan jarak rerata dalam *cluster* serta jarak antara objek dengan *cluster* yang berbeda (Aini dkk., 2014), sesuai dengan definisi perhitungan rata-rata $s(i)$ yakni :

$$SC = \frac{1}{n} \sum_{i=1}^n s(i)$$

Dengan,

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}, \min d(i, C), a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i, j)$$

$b(i)$: nilai terkecil dari jarak rerata objek i dengan semua objek pada *cluster* lain C

$a(i)$: rerata jarak objek ke- i dengan semua objek yang berada di dalam satu *cluster* A

Rentang nilai indeks *silhouette* berkisar dari -1 hingga 1. Pembentukan kelompok dengan kualitas yang baik ditunjukkan dengan indeks *silhouette* yang semakin besar atau mendekati 1.

3. Metodologi

3.1. Data dan Sumber Data

Penelitian ini menggunakan data sekunder yang bersumber dari website BPS Jawa Timur dengan data yang terkumpul terdiri atas 29 kabupaten dan 9 kota mengenai percepatan pemulihan ekonomi yang telah tercantum sebagai pemenuhan pencapaian sasaran utama pembangunan tahun 2022. Daftar nama kabupaten dan kota sebagai unit observasi disajikan pada Gambar 1 berikut.

No	Kab. dan Kota	No	Kab. dan Kota	No	Kab. dan Kota	No	Kab. dan Kota
1.	Kab. Pacitan	11.	Kab. Bondowoso	21.	Kab. Ngawi	31.	Kot. Blitar
2.	Kab. Ponorogo	12.	Kab. Situbondo	22.	Kab. Bojonegoro	32.	Kot. Malang
3.	Kab. Trenggalek	13.	Kab. Probolinggo	23.	Kab. Tuban	33.	Kot. Probolinggo
4.	Kab. Tulungagung	14.	Kab. Pasuruan	24.	Kab. Lamongan	34.	Kot. Pasuruan
5.	Kab. Blitar	15.	Kab. Sidoarjo	25.	Kab. Gresik	35.	Kot. Mojokerto
6.	Kab. Kediri	16.	Kab. Mojokerto	26.	Kab. Bangkalan	36.	Kot. Madiun
7.	Kab. Malang	17.	Kab. Jombang	27.	Kab. Sampang	37.	Kot. Surabaya
8.	Kab. Lumajang	18.	Kab. Nganjuk	28.	Kab. Pamekasan	38.	Kot. Batu
9.	Kab. Jember	19.	Kab. Madiun	29.	Kab. Sumenep		
10.	Kab. Banyuwangi	20.	Kab. Magetan	30.	Kot. Kediri		

Gambar 1. Unit Observasi Kabupaten dan Kota di Jatim

Penelitian ini menggunakan 3 indikator percepatan pemulihan ekonomi setiap kabupaten dan kota di Jawa Timur yang ditampilkan pada Tabel 1 sebagai berikut.

Tabel 1. Variabel Indikator dalam Analisis *Cluster*

Variabel Pemulihan Ekonomi (%)	Keterangan
Pertumbuhan Ekonomi	Peningkatan jumlah produksi barang dan jasa setiap tahunnya ditandai dengan bertambahnya GNP atau GDP.
Tingkat Pengangguran Terbuka	Presentase jumlah pengangguran terhadap jumlah angkatan kerja.
Rasio Gini	Salah satu alat yang mengukur ketimpangan pendapatan dan kekayaan relatif antar penduduk suatu wilayah.

3.2. Metode Analisis

Pendekatan yang diterapkan dalam menganalisis data pada penelitian ini ialah analisis *cluster hierarki*. Pengolahan dan analisis data dilakukan menggunakan perangkat lunak *RStudio*. Berikut merupakan proses analisis dalam penelitian ini.

1. Mengumpulkan data observasi.
2. Menggambarkan karakteristik variabel penelitian secara deskriptif.
3. Melakukan *clustering* kabupaten dan kota di Jawa Timur berdasarkan indikator percepatan pemulihan ekonomi melalui langkah-langkah berikut:
 - a. Melakukan uji kerepresentatifan sampel dan menghindari adanya multikolinieritas sebagai syarat uji asumsi dalam *clustering*.
 - b. Melakukan *clustering* hierarki menggunakan pengukuran jarak *Euclidean* yang diterapkan pada model *single, complete, average, dan ward's linkage*.
 - c. Perbandingan dalam analisis ini dilakukan dengan menggunakan *Internal Clustering Validation* terhadap variasi jumlah *cluster* untuk menentukan jumlah *cluster* optimal dan Koefisien *Cophenetic* untuk menentukan satu model *cluster* hierarki terbaik.
4. Mendeskripsikan karakteristik tiap *cluster* yang telah terbentuk.
5. Merumuskan rekomendasi kebijakan untuk pemerintah guna mempercepat pemulihan ekonomi di Jawa Timur berdasarkan hasil analisis *cluster*.

4. Hasil dan Pembahasan

4.1. Statistika Deskriptif

Secara umum, analisis deskriptif dilakukan untuk mengetahui karakteristik tiap indikator yang digunakan. Pada penelitian ini menampilkan nilai mean, minimum, maksimum, serta *standar deviation* yang disajikan pada Tabel 2.

Tabel 2. Statistika Deskriptif

Indikator Percepatan Pemulihan Ekonomi	Mean	Minimum	Maximum	StDev
Pertumbuhan Ekonomi	4,6332	-6,16 (Kab. Bojonegoro)	8,88 (Kab. Tuban)	2,42995
Tingkat Pengangguran Terbuka	5,2732	1,36 (Kab. Sumenep)	8,80 (Kab. Sidoarjo)	1,77542
Rasio Gini	33,3842	26,60 (Kab. Sumenep)	42,10 (Kota Malang)	3,75518

Berdasarkan Tabel 3 terlihat bahwa secara umum variabel rasio gini memiliki nilai standar deviasi yang relatif tinggi (3,75518), yang berarti menunjukkan variasi lebih besar dalam tingkat ketidaksetaraan pendapatan antar Kabupaten dan Kota. Hal ini menunjukkan bahwa ada variasi yang cukup signifikan dalam distribusi pendapatan di beberapa Kabupaten dan Kota di Jawa Timur. Sebaliknya, variabel tingkat pengangguran terbuka mempunyai nilai standar deviasi yang relatif rendah sebesar 1,77542 yang menunjukkan bahwa variasi tingkat pengangguran antar wilayah tidak terlalu signifikan atau cenderung stabil. Selain itu terlihat bahwa nilai rata – rata variabel rasio gini lebih besar dibandingkan variabel yang lain, meskipun rata – rata variabel Rasio Gini sebesar 33,3842 hal ini dianggap moderat karena nilainya tidak terlalu rendah (menandakan tingkat ketidaksetaraan pendapatan yang rendah) atau terlalu tinggi (menandakan ketidaksetaraan yang sangat tinggi), yang mana menunjukkan bahwa tingkat ketidaksetaraan pendapatan di Jawa Timur masih dalam batas yang wajar.

4.2. Uji Asumsi

Langkah awal dalam analisis *cluster* ialah dilakukannya uji asumsi. Berdasarkan Uji *Kaiser-Mayer-Olkin (KMO)* yang dilakukan pada ketiga variabel, diperoleh nilai 0.5847 dimana nilai KMO tersebut lebih dari 0.5 sehingga sampel memenuhi asumsi representatif dan layak untuk setiap variabel.

Tabel 3 Uji Non Multikolinearitas

Indikator Percepatan Pemulihan Ekonomi	Pertumbuhan Ekonomi	Tingkat Pengangguran Terbuka	Rasio Gini
Pertumbuhan Ekonomi	1,0000	0,2371	0,4393
Tingkat Pengangguran Terbuka	0,2371	1,0000	0,5097
Rasio Gini	0,4393	0,5097	1,0000

Sedangkan berdasarkan matriks korelasi pada Tabel 3, dapat dilihat koefisien korelasi tiap variabel kurang dari 0.8, menunjukkan tidak terjadi multikoleniaritas pada data penelitian sehingga data telah memenuhi asumsi non multikolinearitas yang diperlukan untuk analisis *cluster*.

4.3. Pengelompokan Hierarki

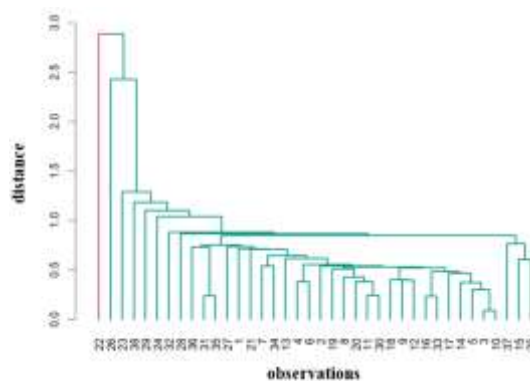
Pengelompokan hierarki pada penelitian ini yakni mengukur jarak Euclidean yang diterapkan pada model *single linkage*, *complete linkage*, *average linkage*, dan *ward's linkage*. Perbandingan dalam analisis ini dilakukan dengan menggunakan *Internal Clustering Validation* terhadap variasi jumlah *cluster* sebanyak 2 hingga 7 *cluster* untuk menentukan jumlah *cluster* optimal pada suatu model *cluster* hierarki.

Single Linkage

Hasil uji validitas internal pada Tabel 4 bahwa ketiga indeks validitas menunjukkan pengelompokan *cluster* optimal pada 2 *cluster*.

Tabel 4. Uji Validitas Internal untuk *Single Linkage*

Indeks Validitas	Jumlah Cluster					
	2	3	4	5	6	7
<i>Connectivity</i>	3,0718	5,8579	8,7869	11,9659	15,8115	19,0405
<i>Dunn</i>	0,5163	0,4338	0,2304	0,2114	0,2283	0,2152
<i>Silhouette</i>	0,5742	0,4219	0,1145	0,0689	-0,0002	-0,0295



Gambar 2. Dendrogram *Single Linkage*

Dua *cluster* optimal yang terbentuk pada model ini disajikan dalam bentuk dendrogram pada Gambar 2 yang masing-masing berisi 1 dan 37 anggota.

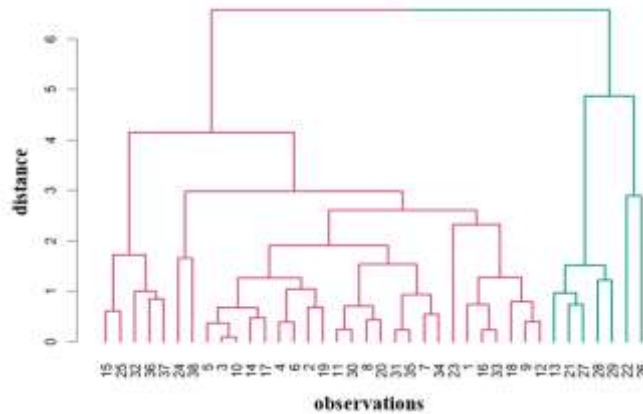
Complete Linkage

Berdasarkan hasil uji validitas internal pada Tabel 5 menunjukkan indeks *dunn* tertinggi ada pada 6 *cluster*, sementara indeks *connectivity* dan *silhouette* menunjuk pada 2 *cluster*, sehingga pada model ini pengelompokan *cluster* optimal terdapat pada *cluster* berjumlah 2.

Tabel 5. Uji Validitas Internal untuk *Complete Linkage*

Indeks Validitas	Jumlah Cluster					
	2	3	4	5	6	7
<i>Connectivity</i>	12,1020	13,2603	20,9825	26,3365	27,4794	37,3278

Indeks Validitas	Jumlah Cluster					
	2	3	4	5	6	7
<i>Dunn</i>	0,1272	0,1494	0,2079	0,2141	0,2375	0,2073
<i>Silhouette</i>	0,4207	0,3906	0,3283	0,3213	0,3114	0,2544



Gambar 3. Dendrogram Complete Linkage

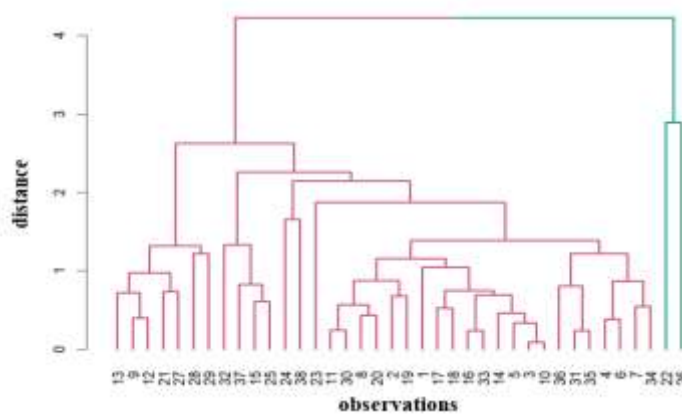
Dua cluster optimal yang terbentuk pada model complete linkage disajikan dalam bentuk dendrogram pada Gambar 3 yang menunjukkan masing-masing berisi 32 dan 6 anggota.

Average Linkage

Hasil uji validitas internal pada model ini ditunjukkan oleh Tabel 6, bahwa indeks dunn tertinggi ada pada 2 atau 3 cluster, sementara indeks connectivity dan silhouette menunjuk pada 2 cluster, sehingga pada model ini pengelompokan cluster optimal terdapat pada cluster berjumlah 2.

Tabel 6. Uji Validitas Internal untuk Average Linkage

Indeks Validitas	Jumlah Cluster					
	2	3	4	5	6	7
<i>Connectivity</i>	4,7151	5,8579	12,6056	19,0155	24,1194	27,0484
<i>Dunn</i>	0,4338	0,4338	0,0995	0,1211	0,1463	0,1463
<i>Silhouette</i>	0,5372	0,4219	0,3780	0,3280	0,3387	0,3211



Gambar 4. Dendrogram Average Linkage

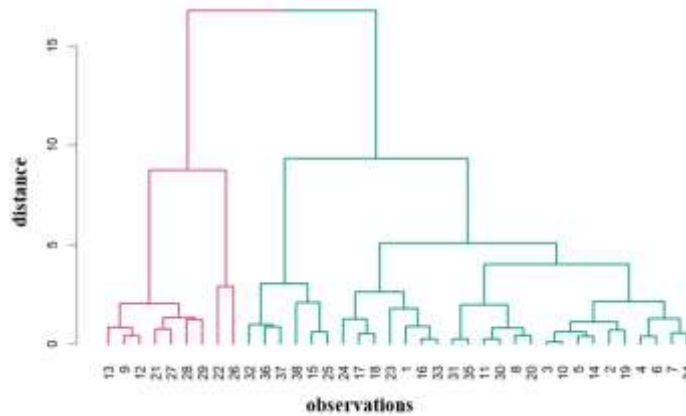
Dua cluster optimal yang terbentuk pada model average linkage disajikan dalam bentuk dendrogram pada Gambar 4 yang masing-masing berisi 36 dan 2 anggota.

Ward's Linkage

Pada model ini, berdasarkan hasil uji validitas internal pada Tabel 7 menunjukkan bahwa indeks dunn tertinggi ada pada 7 cluster, sementara indeks connectivity dan silhouette menunjuk pada 2 cluster, sehingga pada model ini pengelompokan cluster optimal terdapat pada cluster berjumlah 2.

Tabel 7. Uji Validitas Internal untuk Ward's Linkage

Indeks Validitas	Jumlah Cluster					
	2	3	4	5	6	7
Connectivity	9,9948	19,6103	21,0782	31,2873	40,1091	43,9286
Dunn	0,0829	0,0829	0,1328	0,1328	0,1328	0,1396
Silhouette	0,4068	0,3195	0,3513	0,2796	0,2595	0,2715



Gambar 5. Dendrogram Ward's Linkage

Dua cluster optimal yang terbentuk pada model ward's linkage disajikan dalam bentuk dendrogram pada Gambar 5 yang masing-masing berisi 9 dan 29 anggota.

Perbandingan Model Clustering Terbaik

Analisis berikutnya dilakukan dengan perbandingan jarak pengukuran kemiripan menggunakan koefisien korelasi cophenetic yang didapat dari jarak objek dendogram yang telah terbentuk untuk menentukan salah satu model pengelompokan terbaik.

Tabel 8. Koefisien Korelasi Cophenetic

Single linkage	Complete linkage	Average linkage	Ward's linkage
0,7519	0,8970	0,8251	0,9311

Berdasarkan Tabel 8, koefisien korelasi cophenetic tertinggi didapatkan 0,9311, sehingga model terbaik dalam metode clustering hierarki pada penelitian ini adalah Ward's linkage.

4.4. Evaluasi dan Hasil Pengelompokan

Berdasarkan analisis uji Internal Clustering Validation serta koefisien korelasi cophenetic yang dilakukan terhadap keempat model pengelompokan, jumlah cluster optimal untuk clustering kabupaten dan kota di Jawa Timur dilihat dari percepatan pemulihan ekonomi ialah 2 cluster yang diperoleh dari perbandingan indeks connectivity dan indeks silhouette terhadap masing-masing model dimana hasil uji sama-sama menunjukkan optimal pada 2 cluster, sedangkan berdasarkan koefisien cophenetic diperoleh model terbaik dalam metode clustering hierarki pada penelitian ini adalah Ward's linkage.

Tabel 9. Hasil Clustering berdasarkan Model Ward's Linkage

Kabupaten dan Kota Jawa Timur	Cluster	
	1	2
Kabupaten	Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Malang, Lumajang, Banyuwangi, Bondowoso, Pasuruan, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Tuban, Lamongan, Gresik.	Jember, Situbondo, Probolinggo, Ngawi, Bojonegoro, Bangkalan, Sampang, Pamekasan, Sumenep.
Kota	Kediri, Blitar, Malang, Probolinggo, Pasuruan, Mojokerto, Madiun, Surabaya, Batu.	

4.5. Karakteristik Tiap Cluster yang Terbentuk

Pemetaan hasil clustering pada Gambar 6 menunjukkan terdapat 2 warna yang digunakan untuk menunjukkan hasil pengelompokan kabupaten dan Kota di Provinsi Jawa Timur dilihat dari percepatan

pemulihan ekonomi menggunakan model *Ward's*. Wilayah berwarna kuning menunjukkan *cluster* 1 yang terbentuk yaitu sebanyak 29 kabupaten dan kota, warna hijau untuk *cluster* 2 sebanyak 9 kabupaten.



Gambar 6. Peta Cluster Kab. dan Kota dengan Percepatan Pemulihan Ekonomi di Jatim

Tabel 10. Profilisasi Cluster

Cluster	Kategori Rata-Rata Indikator	Pertumbuhan Ekonomi (%)	Tingkat Pengangguran Terbuka (%)	Rasio Gini (%)
1	Tinggi	5,40	5,81	34,7
2	Rendah	2,16	3,53	2,0

Berdasarkan hasil *clustering* kabupaten dan kota di Jawa Timur dengan model *Ward's Linkage*, dapat dilihat pola karakteristik untuk tiap *cluster* yang telah terbentuk pada Tabel 11 berikut.

Tabel 11. Karakteristik Tiap Cluster

Cluster	Kab. dan Kota Jatim	Karakteristik
1	Pacitan, Ponorogo, Trenggalek, Tulungagung, Blitar, Kediri, Malang, Lumajang, Banyuwangi, Bondowoso, Pasuruan, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Tuban, Lamongan, Gresik, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota Batu.	<ul style="list-style-type: none"> a. Percepatan pemulihan ekonomi yang tinggi dengan rata-rata 5,40 %. b. Tingkat pengangguran terbuka tinggi dengan rata-rata 5,81 % menunjukkan tingkat penyerapan tenaga kerja yang kurang optimal di pasar kerja sehingga peluang kerja lebih sedikit. c. Rasio gini tinggi dengan rata-rata 34,7 % menunjukkan bahwa distribusi pendapatan belum merata sebaik <i>cluster</i> 2 sehingga pada <i>cluster</i> ini kesenjangan ekonomi masih cenderung tinggi. d. Untuk mempercepat pemulihan ekonomi pada <i>cluster</i> ini, indikator tingkat pengangguran terbuka dan rasio gini masih perlu untuk diminimalisasikan.
2	Jember, Situbondo, Probolinggo, Ngawi, Bojonegoro, Bangkalan, Sampang, Pamekasan, Sumenep.	<ul style="list-style-type: none"> a. Percepatan pemulihan ekonomi yang rendah dengan rata-rata 2,16 %. b. Tingkat pengangguran terbuka tinggi dengan rata-rata 3,53 % menunjukkan tingkat penyerapan tenaga kerja cukup optimal di pasar kerja sehingga peluang kerja lebih banyak atau kebijakan ketenagakerjaan yang lebih efektif. c. Rasio gini tinggi dengan rata-rata 29 % menunjukkan adanya pemerataan yang lebih baik dalam distribusi pendapatan sehingga tingkat kesenjangan ekonomi lebih rendah dibandingkan dengan <i>cluster</i> 1. d. Untuk mempercepat pemulihan ekonomi pada <i>cluster</i> ini, indikator tingkat pertumbuhan ekonomi perlu ditingkatkan lagi.

Clustering ini diharapkan dapat memberikan pemahaman yang baik terkait pola pemulihan ekonomi di Jawa Timur dan membantu dalam merencanakan kebijakan dan strategi yang lebih efektif untuk mempercepat pemulihan ekonomi di setiap *cluster*. Kebijakan yang ditujukan untuk *Cluster 2* dapat difokuskan pada upaya pemulihan ekonomi yang lebih lanjut, seperti stimulus ekonomi atau insentif investasi. Sementara itu, kebijakan untuk *Cluster 1* dapat difokuskan pada peningkatan pemerataan pendapatan melalui program-program redistribusi atau peningkatan akses ke peluang kerja. Dengan mempertimbangkan kebutuhan dan potensi masing-masing *cluster*, kebijakan yang diimplementasikan dapat lebih efisien dan berdampak positif dalam memulihkan ekonomi secara keseluruhan.

5. Kesimpulan

Penelitian menunjukkan bahwa model terbaik dalam metode *clustering* hierarki ini adalah *Ward's linkage* dengan koefisien korelasi *cophenetic* tertinggi sebesar 0,9311. Didapatkan dua *cluster* optimal dalam *clustering* kabupaten dan kota di Jawa Timur berdasarkan percepatan pemulihan ekonomi. *Cluster 1* menunjukkan pemulihan ekonomi yang lebih signifikan dengan pertumbuhan ekonomi yang lebih tinggi, menandakan adanya aktivitas ekonomi yang lebih kuat dan pemulihan yang lebih cepat. Di sisi lain, *Cluster 2* menunjukkan tingkat penyerapan tenaga kerja yang lebih optimal di pasar kerja, mengindikasikan adanya peluang kerja yang lebih banyak atau kebijakan ketenagakerjaan yang lebih efektif serta memiliki pemerataan pendapatan yang lebih merata dengan rasio gini yang lebih rendah. Perbedaan karakteristik antara kedua *cluster* ini dapat menjadi referensi pemerintah dalam optimalisasi strategi pengembangan sesuai dengan kebutuhan masing-masing *cluster*, misalnya program stimulus ekonomi pada *cluster 2* perlu dioptimalkan lagi sedangkan bagi *cluster 1* diharapkan dapat mengoptimalkan redistribusi pendapatan serta akses peluang kerja.

6. Daftar Pustaka

- Afifi, A., May, S., Donatello, R., & Clark, V.A. (2019). *Practical Multivariate Analysis* (6th ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315203737>.
- Aini, F. N., Palgunadi, S., & Anggrainingsih, R. (2014). *Clustering Business Process Model Petri Net*. *Jurnal Itsmart*, 3(2), 47–51.
- Bezdek, J. C., & Pal, N. R. (1995). Cluster validation with generalized Dunn's indices. *Proceedings - 1995 2nd New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, ANNES 1995, March*, 190–193. <https://doi.org/10.1109/ANNES.1995.499469>
- Badan Pusat Statistik Jawa Timur. (2023). *Tingkat Pengangguran Terbuka (TPT) Menurut Kabupaten/Kota (Persen), 2021-2023*. Jawa Timur: Badan Pusat Statistik.
- Badan Pusat Statistik Jawa Timur. (2022). *Gini Rasio Menurut Kabupaten/Kota di Jawa Timur 2020-2022*. Jawa Timur: Badan Pusat Statistik.
- Badan Pusat Statistik Jawa Timur. (2022). *Pertumbuhan Ekonomi Menurut Kabupaten/Kota (Persen) 2018-2022*. Jawa Timur: Badan Pusat Statistik.
- Da Silva, A. R., & Dias, C. T. dos S. (2013). A cophenetic correlation coefficient for tocher's method. *Pesquisa Agropecuaria Brasileira*, 48(6), 589–596. <https://doi.org/10.1590/S0100-204X2013000600003>.
- Hair, J., Anderson, R., Babin, B., & Black, W. (2010). *Multivariate Data Analysis*.pdf. In *Australia : Cengage: Vol. 7 edition* (p. 758).
- Halim, N. N., & Widodo, E. (2017). Clustering Dampak Gempa Bumi di Indonesia Menggunakan Kohonen Self Organizing Maps (SOM). *Prosiding SI MaNIs (Seminar Nasional Integrasi Matematika Dan Nilai-Nilai Islami)*, 1(1): 188–194.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*.: Pearson Prentice Hall. In *Pearson Prentice Hall*.
- Naes, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *Multivariate calibration and classification*. West Sussex: NIR Publication.
- Nahdliyah, M. A., Widiharih, T., & Prahutama, A. (2019). Metode K-Medoids *Clustering* dengan Validasi Silhouette Index dan C-Index (Studi Kasus Jumlah Kriminalitas Kabupaten dan kota di Jawa Tengah Tahun 2018). *Jurnal Gaussian*, 8(2), 161-170. <https://doi.org/10.14710/j.gauss.8.2.161-170>.
- Prasetyo, E. (2012). *Data Mining: Konsep dan Aplikasi Menggunakan Matlab*. Yogyakarta: Andi Offset.
- Saracli, S., Dogan, N., dan Dogan, I. (2013). *Comparison of Hierarchical Cluster Analysis Methods by Cophenetic Correlation*. *Journal of Inequalities and Applications*, Hal. 203. <https://doi.org/10.1186/1029-242X-2013-203>.
- Widarjono, A. (2010). *Analisis Statistika Multivariat Terapan*. Edisi pertama. Yogyakarta: UPP STIM YKPN
- Wierzchoń, S. T., & Kłopotek, M. A. (2018). *Cluster Analysis*. In *Studies in Big Data* (Vol. 34). https://doi.org/10.1007/978-3-319-69308-8_2.