

## ANALISIS TINGKAT KEMISKINAN DENGAN ALGORITMA K-MEANS MENGGUNAKAN RAPIDMINER DI TINGKAT KOTA KABUPATEN DI JAWA TENGAH

Iqbal Tawakal<sup>1)</sup>, M.Makmun Effendi<sup>2)</sup>, Annisa Maulana Majid<sup>3)</sup>

<sup>1, 2, 3)</sup> Teknik Informatika Universitas Pelita Bangsa

email : iqbaltauwakal733@mhs.pelitabangsa.ac.id<sup>1)</sup>, effendiyan@pelitabangsa.ac.id<sup>2)</sup>, annisa.maulanamajid@pelitabangsa.ac.id<sup>3)</sup>

### Abstraksi

Permasalahan kemiskinan di Provinsi Jawa Tengah masih menjadi isu krusial yang membutuhkan penanganan berbasis data yang akurat dan terstruktur. Ketiadaan pemetaan yang tepat sering kali menyebabkan distribusi bantuan sosial tidak efektif dan tidak menyentuh kelompok masyarakat yang paling membutuhkan. Penelitian ini bertujuan untuk mengidentifikasi pola kemiskinan antar wilayah kota atau kabupaten di Jawa Tengah dengan memanfaatkan algoritma K-Means melalui platform RapidMiner. Analisis dilakukan terhadap data sekunder dari Badan Pusat Statistik periode 2020–2024, mencakup berbagai indikator seperti garis kemiskinan, persentase penduduk miskin, tingkat pengangguran, serta akses terhadap air bersih, sanitasi, dan bantuan sosial. Proses analisis mengikuti tahapan *Knowledge Discovery in Databases*, dari pemilihan data hingga evaluasi hasil. Penelitian ini menghasilkan lima kelompok kemiskinan dengan akurasi klaster tinggi berdasarkan nilai Davies-Bouldin Index sebesar 0,009. Hasil pengelompokan ini memberikan gambaran yang lebih objektif tentang distribusi kemiskinan di Jawa Tengah dan dapat dimanfaatkan sebagai dasar perumusan kebijakan yang lebih tepat sasaran dan efisien dalam upaya pengentasan kemiskinan.

### Kata Kunci :

Kemiskinan, Algoritma K-Means, RapidMiner, Jawa Tengah, Clustering, Davies-Bouldin Index.

### Abstract

*The problem of poverty in Central Java Province is still a crucial issue that requires handling based on accurate and structured data. The absence of proper mapping often causes the distribution of social assistance to be ineffective and not reach the community groups that need it most. This study aims to identify poverty patterns between cities or regencies in Central Java by utilizing the K-Means algorithm through the RapidMiner platform. The analysis was carried out on secondary data from the Central Statistics Agency for the period 2020–2024, covering various indicators such as the poverty line, percentage of poor people, unemployment rate, and access to clean water, sanitation, and social assistance. The analysis process follows the stages of Knowledge Discovery in Databases, from data selection to evaluation of results. This study produced five poverty groups with high cluster accuracy based on the Davies-Bouldin Index value of 0.009. The results of this grouping provide a more objective picture of the distribution of poverty in Central Java and can be used as a basis for formulating more targeted and efficient policies in poverty alleviation efforts.*

### Keywords :

Poverty, K-Means Algorithm, RapidMiner, Central Java, Clustering, Davies-Bouldin Index.

### Pendahuluan

Kemiskinan masih menjadi tantangan utama pembangunan di Indonesia, termasuk di Provinsi Jawa Tengah. Data Badan Pusat Statistik (BPS) tahun 2023 menunjukkan bahwa tingkat kemiskinan di Jawa Tengah mencapai 10,98%, menempatkannya sebagai provinsi dengan jumlah penduduk miskin tertinggi ke-10 di Indonesia. Meskipun pemerintah telah melaksanakan berbagai program bantuan sosial, distribusi bantuan tersebut kerap tidak merata dan kurang tepat sasaran. Salah satu penyebabnya adalah kurangnya validasi dan analisis data yang akurat, sehingga informasi tentang daerah yang benar-benar membutuhkan bantuan tidak dapat terpetakan dengan baik.

Penelitian ini bertujuan untuk mengelompokkan kotaataukabupaten di Jawa Tengah berdasarkan indikator-indikator kemiskinan menggunakan algoritma K-Means melalui perangkat lunak RapidMiner. Dengan analisis ini, diharapkan pola distribusi kemiskinan dapat terlihat secara lebih objektif dan dapat menjadi acuan dalam menyusun kebijakan pengentasan kemiskinan yang lebih tepat sasaran.

Penelitian ini dibatasi hanya pada wilayah Provinsi Jawa Tengah dengan rentang data tahun 2020 hingga 2024. Indikator yang digunakan meliputi garis kemiskinan, jumlah dan persentase penduduk miskin, tingkat pengangguran, akses terhadap air bersih dan sanitasi, serta penerima bantuan sosial. Algoritma K-Means digunakan sebagai metode utama dalam

pengelompokan data, dengan evaluasi kualitas klaster menggunakan indeks Davies-Bouldin. RapidMiner digunakan sebagai alat bantu utama dalam proses analisis dan visualisasi data.

### **Tinjauan Pustaka**

Berbagai penelitian sebelumnya telah mengkaji penerapan algoritma *K-Means* dalam mengelompokkan wilayah berdasarkan tingkat kemiskinan. Assyifa Khalif dkk. (2024) mengelompokkan 34 provinsi di Indonesia menggunakan 12 indikator sosial-ekonomi, dengan hasil tiga klaster utama dan *Silhouette Score* sebesar 0,7416 [1]. Nugraha (2022) memetakan 27 kabupaten/kota di Jawa Barat ke dalam tiga kategori kemiskinan dengan evaluasi skor 0,55 [2]. Penelitian oleh Mayasari dan Nugraha (2023) di Jawa Tengah membentuk tiga klaster berdasarkan jumlah penduduk miskin, pertumbuhan penduduk, dan tingkat pengangguran, meskipun tanpa evaluasi eksplisit [3]. Di sisi lain, Fuji Astri Martanto (2024) dan Irfan & Lut Faizal (2024) menggunakan *Davies-Bouldin Index* (DBI) dalam evaluasi klasterisasi di Jawa Barat dan Sulawesi Selatan, dengan nilai DBI masing-masing sebesar 0,004 dan 0,398 [4], [5]. Secara umum, pendekatan yang digunakan bervariasi, dengan mayoritas penelitian masih terbatas pada indikator dasar dan tidak menyebutkan penggunaan alat bantu perangkat lunak secara eksplisit.

Penelitian ini berbeda dari sebelumnya karena memanfaatkan *RapidMiner* dalam proses pengelompokan, serta menggabungkan indikator kemiskinan yang lebih luas, seperti akses air minum dan sanitasi layak, serta penerima bantuan sosial. Evaluasi menggunakan *Davies-Bouldin Index* menghasilkan nilai 0,009, menunjukkan kualitas klasterisasi yang sangat baik. Selain itu, fokus wilayah yang terbatas pada Provinsi Jawa Tengah memberikan kedalaman analisis spasial yang lebih tinggi, terutama dengan visualisasi data melalui peta dan grafik. Hasil penelitian ini diharapkan dapat menjadi acuan kebijakan yang lebih tepat sasaran dalam upaya pengentasan kemiskinan di tingkat daerah.

### **Kemiskinan**

Kemiskinan adalah istilah yang digunakan untuk menggambarkan seseorang yang tidak dapat memenuhi kebutuhan dasar mereka [6]. Rendahnya pendapatan yang diperlukan untuk memenuhi kebutuhan pokok seperti makanan, pakaian, dan papan menunjukkan ketidakmampuan ekonomi untuk memenuhi standar hidup rata-rata masyarakat setempat. Kemiskinan tidak hanya dilihat dari satu dimensi saja seperti konsumsi atau pendapatan. Kemiskinan mencakup dimensi ekonomi, sosial, dan akses terhadap kebutuhan dasar, sehingga tidak hanya mencerminkan ketidakmampuan pendapatan tetapi juga kualitas hidup secara menyeluruh [7].

### **Rapidminer**

Suatu platform perangkat lunak pembelajaran mesin dan ilmu data yang sangat kuat, Banyak alat tersedia untuk pemodelan, evaluasi, implementasi, dan persiapan data. Meskipun dirancang untuk menjadi mudah digunakan, Rapidminer memungkinkan orang tanpa pengalaman pemrograman untuk membuat dan menguji berbagai model dengan cepat. Memanfaatkan antarmuka drag-and-drop, pengguna dapat membuat alur kerja untuk memproses dan menganalisis data. Salah satu keunggulan perangkat lunak ini adalah kemampuan untuk bekerja dengan berbagai sumber data, seperti basis data. Selain itu, RapidMiner menyertakan berbagai operator bawaan yang mencakup semua tahapan proses pemrosesan data, seperti pemodelan, pembersihan data, dan pemilihan fitur [4].

### **QGIS**

Sistem yang berfungsi untuk mengumpulkan, menyimpan, mengubah, menganalisa, mengatur, dan menampilkan berbagai jenis data geografis dikenal sebagai sistem informasi geografis (SIG). *Quantum GIS* adalah perangkat lunak *open source* yang gratis dan mudah digunakan yang berjalan di berbagai platform seperti *Linux*, *Unix*, *Mac OS X*, dan *Windows*. Dengan menggunakan QGIS, data peta dapat ditampilkan dan dibuat dalam format shapefile, geotiff, dan format lainnya. Selain itu, perangkat SIG ini mendukung plugin untuk melakukan tugas tertentu, seperti menampilkan pengelompokan daerah berdasarkan data [8].

### **Data Mining**

Data mining adalah proses analisis data yang bertujuan untuk mengekstraksi pola atau informasi penting dari kumpulan data kompleks. Data mining menggunakan algoritma seperti regresi, klasifikasi, pengelompokan, dan asosiasi untuk membantu membuat keputusan yang lebih baik dan mengoptimalkan penggunaan sumber setiap hari. Untuk menemukan pola yang tidak terlihat secara langsung, memprediksi hasil yang akan datang, membagi data menjadi kelompok, dan menemukan anomali, proses ini memanfaatkan kecerdasan buatan, matematika, dan statistika [9].

### **Clustering**

Clustering adalah teknik dalam analisis data yang mengelompokkan objek atau data ke dalam kelompok atau kluster berdasarkan karakteristik yang sebanding. Tujuan utama clustering adalah kelompok yang memiliki banyak kesamaan internal tetapi juga sedikit perbedaan. Tingkat perbedaan yang tinggi antara satu kluster dan tingkat keseragaman di antara anggota kluster merupakan syarat jika kluster bisa dikatakan baik [10].

### K.Means

Teknik klasterisasi yang dikenal sebagai algoritma K-means mengelompokkan data berdasarkan centroid atau pusat klaster terdekat. Algoritma ini menghitung tingkat kemiripan klaster berdasarkan jarak terkecil antara centroid dan data. Tujuan algoritma ini adalah untuk mengelompokkan data dengan tingkat kesamaan tertinggi di antara kluster [4]. Proses implementasi algoritma K-Means adalah sebagai berikut:

- Menghitung jumlah cluster (k).
- Menetapkan pusat clusteratacentroid secara acak dan sebanyak jumlah cluster.
- Untuk mengetahui jarak antara data dan pusat cluster, gunakan rumus persamaan *Euclidean Distance* :

$$D_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + \dots + (n_i - n_j)^2}$$

Keterangan :

$D_{ij}$  = Jarak antara i dan j

$x_i$  = Koordinat x objek

$x_j$  = Koordinat x pusat

$y_i$  = Koordinat y objek

$y_j$  = Koordinat y pusat

- Untuk menemukan centroid baru, hitung rata-rata hasil data pada cluster dengan posisi yang sama.
- Menghitung kembali data dari pusat cluster yang baru saja diperoleh. Jika ada perubahan pada pusat cluster, proses diulangi dari Langkah ketiga, tetapi jika tidak ada perubahan, proses pencarian dihentikan.

### Davies Bouldin Index

*Davies Boludin Index* ialah sistem untuk menilai hasil clustering yang telah dibuat. *Index Davies-Bouldin* menghasilkan proses evaluasi klaster internal. Hasil clustering dianggap lebih baik jika DBI yang dihasilkan mendekati 0. Tujuan penggunaan indeks ini adalah untuk memperpendek jarak antara titik-titik dalam klaster sambil memperbesar jarak antara kluster [11].

### Metode Penelitian

Di penelitian ini, teknik analisis “*Knowledge Discovery in Databases*” digunakan. *Knowledge Discovery in Databases* adalah sekumpulan prosedur yang digunakan untuk mengeksplorasi dan menganalisis sejumlah besar data untuk menghasilkan informasi dan pengetahuan yang berguna [6]. Proses *Knowledge Discovery in Databases* biasanya terdiri dari langkah-langkah berikut:

#### 1. Data Selection

Data sekunder digunakan dalam penelitian ini. Data ini berasal dari website resmi Badan Pusat

Statistik Jawa Tengah yang mengumpulkan data tentang kemiskinan di Jawa Tengah. Dataset ini berisi data kemiskinan untuk setiap Kabupaten atau Kota di Provinsi Jawa Tengah dari tahun 2020 hingga 2024. Data awal yang dikumpulkan harus dikumpulkan dalam satu file terlebih dahulu karena data yang diterima masih dalam bentuk file terpisah. Selain itu, proses ini mengubah nama indikator menjadi kode abjad A-K2, dengan setiap baris menunjukkan kabupaten atau kota, dan setiap kolom menunjukkan atribut data, atau variabel. Data yang digunakan dalam analisis, berupa indikator kemiskinan dari berbagai kabupaten atau kota di Jawa Tengah dapat dilihat pada Tabel 1 dan keterangan kode indikator yang digunakan dalam dataset dapat dilihat pada Tabel 2.

Tabel 1. Dataset yang digunakan

Kabupaten atau Kota	A	A1	A2	....	K2
Kabupaten	351	363	384	....	92,72
Cilacap	735	367	955		
Kabupaten	406	417	441	....	93,38
Banyumas	250	086	520		
Kabupaten	375	384	407	....	95,5
Purbalingga	199	183	849		
Kabupaten	318	328	351	....	95,55
Banjarnegara	334	679	333		
Kabupaten	380	390	416	....	94,98
Kebumen	557	599	004		
Kabupaten	364	376	393	....	89,27
Purworejo	289	127	731		
Kabupaten	362	373	399	....	96,01
Wonosobo	683	474	180		
Kabupaten	342	353	377	....	93,7
Magelang	430	608	497		
.....	....	....	....	....	....
.....	....	....	....	....	....
Kota	460	480	513	....	67,23
Pekalongan	789	415	243		
Kota Tegal	502	523	565	....	66,62
	031	413	826		

Tabel 2. Keterangan Dataset

Kode	Indikator	Tahun	Ket
A – A4	Garis Kemiskinan	2020 – 2024	Rp atau kapita aubln
B – B4	Jumlah Penduduk Miskin	2020 – 2024	Ribu Jiwa
C – C4	Persentase Penduduk Miskin	2020 – 2024	Per森
D – D4	Tingkat Pengangguran Terbuka	2020 – 2024	Per森
E – E4	Indeks Kedalaman Kemiskinan (P1)	2020 – 2024	Per森

F – F4	Indeks Keparahan Kemiskinan (P2)	2020 – 2024	Persen
G – G3	Persentase Rumah Tangga Yang Memiliki Akses Terhadap Sanitasi Layak	2020 – 2023	Persen
H – H3	Persentase Rumah Tangga Yang Memiliki Akses Terhadap Air Minum Layak	2020 – 2023	Persen
I – I3	Persentase Penerima Bansos Pangan	2020 – 2023	Persen
J – J3	Persentase Penerima PKH	2020 – 2023	Persen
K – K2	Persentase Rumah Tangga Dengan Status Pengusahaan Bangunan Tempat Tinggal Milik Sendiri	2020, 2021, 2023	Persen

## 2. Data Pre-Processing

Pre-processing data yang dimaksud adalah membersihkan data mulai dari menghapus data yang tidak lengkap, terdapat data error, terdapat ketidakkonsistensi data [12]. Pada tahap preprocessing data, dilakukan pengecekan untuk memastikan apakah ada nilai yang hilang atau hilang dalam kumpulan data. Penelitian ini menggunakan data lengkap, dan hasil statistik menunjukkan bahwa tidak ada nilai yang hilang.

## 3. Data Transformation

Tahap ini digunakan untuk mengubah data ke dalam format yang diperlukan untuk proses data mining. Karena variabel dalam penelitian terdiri dari skala yang berbeda, seperti jumlah penduduk (kontinu) dan persentase (proporsi). Selanjutnya, data dinormalisasi untuk memastikan skala yang konsisten.

## 4. Data Mining

Proses ini dilakukan untuk menghasilkan informasi yang bermanfaat atau bernilai. Data mining yang diterapkan adalah K-Means Clustering. Langkah pertama melibatkan pemilihan pusat setiap klaster secara acak. Jarak antara setiap titik data individu dan pusat klaster kemudian dihitung. Pengelompokan mencapai stabilitas kluster setelah berulang kali.

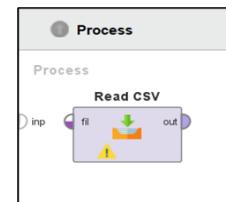
## 5. Evaluation dan Visualisation

Tujuan dari evaluasi pengelompokan adalah untuk mengetahui seberapa baik hasil pengelompokan tersebut. Metode evaluasi internal *clustering DBI* digunakan dalam penelitian ini untuk menemukan jumlah cluster yang ideal. Jumlah cluster yang ideal dan kualitas clustering yang lebih baik ditunjukkan oleh nilai DBI yang lebih rendah. Untuk membuat data lebih mudah dipahami dan menarik, tahapan visualisasi ini digunakan. Untuk menunjukkan jumlah dan jarak dari setiap daerah data pada masing-masing cluster yang terbentuk, visualisasi data ini dapat berupa tabel, diagram, dll [13]. Di penelitian ini, penulis menggunakan peta sebagai hasil akhir dalam memvisualisasikan hasil clustering tingkat kemiskinan ditingkat kotaataukabupaten di Jawa Tengah.

## Hasil dan Pembahasan

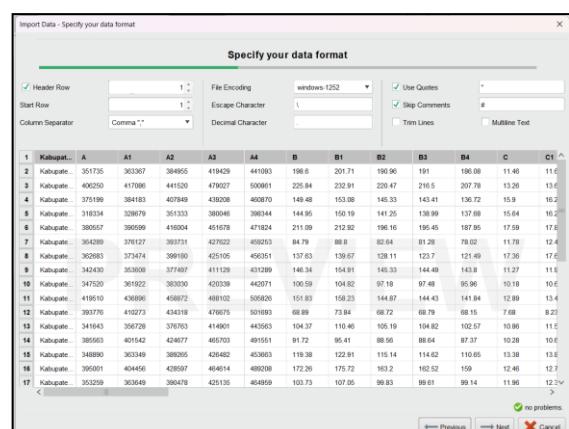
### 1. Data Selection

Langkah pertama yaitu dengan memilih data untuk analisis. Ini dilakukan dengan memilih operator "Read CSV" dalam RapidMiner seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Operator Read CSV

Setelah memilih operator, pilih "Import Data" lalu dataset tersebut akan dimasukkan ke dalam proses pengolahan data RapidMiner. Sebagaimana diperlihatkan dalam Gambar 2.



Gambar 2. Import Dataset

## 2. Data PreProcessing

Pada tahap preprocessing data, nilai yang hilang (*missing values*) dicari dalam dataset. Untuk melakukan ini, proses "Read CSV" dijalankan untuk mengakses dataset yang telah diimpor sebelumnya. Memeriksa komponen "Value Missing" dalam menu statistik perangkat lunak setelah datasetnya terbuka.

Setelah mengecek hasil statistik data, ditunjukkan pada Gambar 3 bahwa tidak ada nilai yang hilang.

Name	Type	Missing
KabupatenKota	Nominal	0
A	Integer	0
A1	Integer	0
A2	Integer	0
A3	Integer	0
A4	Integer	0
B	Real	0
B1	Real	0
B2	Real	0
B3	Real	0

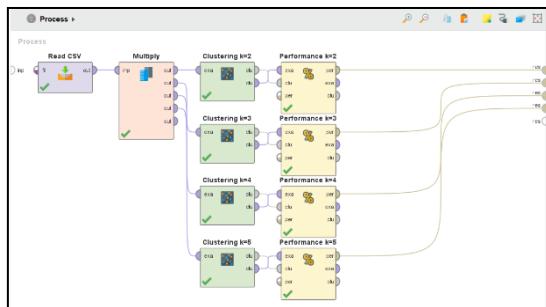
Gambar 3. Missing Values

### 3. Data Transformation

Pada langkah berikutnya, penulis menetapkan peran setiap atribut yang akan digunakan dalam proses analisis. Dengan menggunakan operator ini, penulis dapat menentukan atribut yang akan berfungsi sebagai pengenal unik atau variabel lain tergantung pada kebutuhan analisisnya. Penelitian ini menetapkan atribut "KotaKabupaten" sebagai identitas, yang menunjukkan bahwa atribut ini akan berfungsi sebagai pengenal unik dalam kumpulan data. Pengaturan peran atribut sangat penting untuk memastikan bahwa data dapat diolah dengan benar pada tahapan analisis berikutnya.

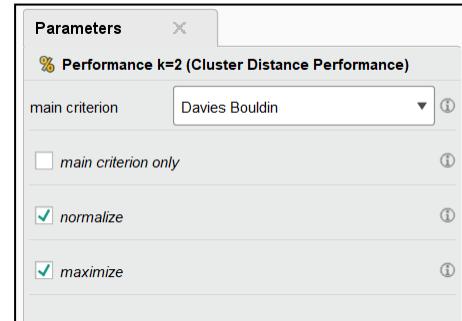
### 4. Data Mining

Operator "Performance" digunakan untuk menemukan nilai K yang paling ideal selama proses klasterisasi menggunakan K-means untuk menghitung nilai DBI. Nilai K berkisar antara 2 sampai 5. Beberapa parameter penting yang digunakan adalah "Max runs 10", "Measure Types: numerical measure", "Numerical measure: Euclidean Distance", dan "Max Step Optimization 100". Pengaturan ini dimaksudkan untuk mengidentifikasi pola klaster yang paling ideal dan memaksimalkan kinerja algoritma. Pemilihan jenis measure type, khususnya Euclidean Distance, sangat penting untuk menentukan kedekatan antar-poin dalam ruang numerik, yang memungkinkan pembentukan klaster yang sesuai dengan karakteristik numerik data kemiskinan. Ilustrasi proses tersebut ditampilkan pada Gambar 4.



Gambar 4. Proses K-Means

Selain yang disebutkan diatas, ada lagi parameter yang sangat penting yaitu Normalize dan Maximize. Tujuan *normalize* adalah untuk memasukkan semua nilai metrik ke dalam rentang yang sama, biasanya antara 0 dan 1, sehingga menjadi lebih mudah dan objektif untuk membandingkan nilai metrik terlepas dari skala awal nilai-nilai dalam dataset. Sedangkan *maximize* digunakan karena hasil akhirnya tidak dikalikan dengan minus satu sehingga menghasilkan nilai DBI positif, Sebagaimana diperlihatkan dalam Gambar 5.



Gambar 5. Parameter Performance

### 5. Evaluation and Visualisation

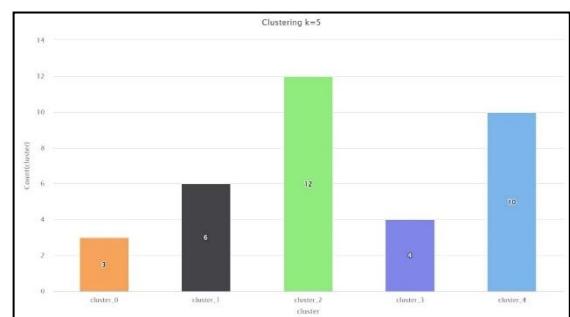
Setelah melakukan pengujian, penulis mendapatkan hasil dari masing – masing jumlah klaster sebagaimana dirangkum dalam Tabel 3.

Tabel 3. Hasil Pengujian Beberapa Nilai K

Jumlah Klaster	Nilai DBI
K=2	0,012
K=3	0,010
K=4	0,010
K=5	0,009

Penelitian ini menggunakan lima klaster karena hasil evaluasi menunjukkan bahwa lima klaster memberikan hasil terbaik, dengan nilai DBI terendah mencapai K = 5 dan nilai 0,009 yang hampir nol. Cluster diberi nama berdasarkan tingkat kemiskinannya: Sangat miskin (Cluster 0), Miskin (Cluster 1), Rentan Miskin (Cluster 2), Sedang (Cluster 3), dan Sejahtera (Cluster 4).

Hasil penelitian clustering ditunjukkan pada Gambar 6 di bawah ini. Dalam grafik ini, sumbu X menunjukkan kelima klaster yang terbentuk, dan sumbu Y menunjukkan jumlah data yang termasuk dalam masing-masing klaster.



Gambar 6. Grafik Bar

Untuk memberikan gambaran yang lebih baik, setiap kabupaten atau kota dimasukkan ke dalam masing-masing klaster sebagai berikut :

#### 1. Cluster 0 (Sangat Miskin)

Wilayah yang tergabung dalam klaster ini menunjukkan tingkat kesejahteraan yang paling rendah, mencerminkan kondisi kemiskinan ekstrem. Kabupaten Banjarnegara, Temanggung, dan Batang menjadi anggota kelompok ini karena memiliki proporsi penduduk miskin yang sangat tinggi, disertai dengan nilai kedalaman dan keparahan kemiskinan yang mencolok dibandingkan wilayah lain. Akses masyarakat terhadap kebutuhan dasar seperti air bersih dan sanitasi sangat terbatas, dan sebagian besar rumah tangga sangat bergantung pada bantuan sosial, baik berupa bantuan pangan maupun Program Keluarga Harapan (PKH), yang menandakan rendahnya kemandirian ekonomi di daerah-daerah tersebut.

#### 2. Cluster 1 (Miskin)

Cluster ini terdiri dari wilayah yang masih menghadapi permasalahan kemiskinan, meskipun tidak separah kelompok sebelumnya. Enam daerah termasuk dalam kategori ini, yaitu Kabupaten Pati, Kudus, Demak, Brebes, serta Kota Salatiga dan Kota Pekalongan. Wilayah-wilayah ini menunjukkan jumlah penduduk miskin yang relatif besar, namun indikator kedalaman dan keparahan kemiskinannya berada pada level menengah. Akses ke infrastruktur dasar seperti sanitasi dan air bersih masih belum memadai, dan tingkat ketergantungan masyarakat terhadap bantuan sosial tetap tinggi, meskipun sedikit lebih baik dibandingkan dengan klaster sangat miskin.

#### 3. Cluster 2 (Rental Miskin)

Daerah yang masuk dalam klaster ini tidak termasuk dalam kategori miskin ekstrem, namun menunjukkan tingkat kerentanan ekonomi yang tinggi, sehingga berisiko kembali terjerumus ke dalam kemiskinan apabila tidak ditangani secara berkelanjutan. Terdiri dari 12 kabupaten, antara lain Banyumas, Klaten, Sukoharjo, Karanganyar, Grobogan, Rembang, Jepara, Semarang, Kendal, Pekalongan, Pemalang, dan Tegal. Wilayah-wilayah ini mencerminkan kondisi ekonomi yang mulai membaik ditinjau dari peningkatan akses terhadap fasilitas dasar dan berkurangnya ketergantungan terhadap bantuan sosial. Namun, stabilitas ekonomi masih rapuh, sehingga perlu intervensi strategis untuk menjaga tren positif ini.

#### 4. Cluster 3 (Sedang)

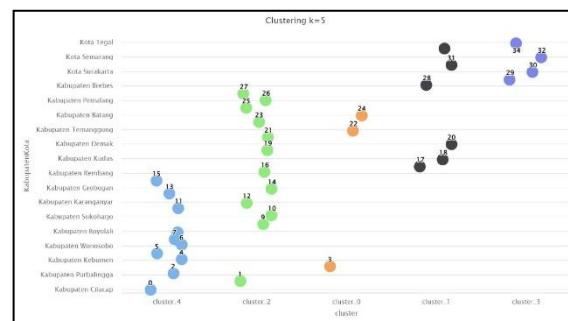
Cluster ini menggambarkan wilayah dengan kondisi sosial-ekonomi yang relatif stabil dan cukup seimbang. Kota Magelang, Surakarta, Semarang, dan Tegal termasuk dalam kelompok ini. Jumlah dan persentase penduduk miskin di wilayah-wilayah ini tergolong rendah, dan hampir seluruh rumah tangga

telah memiliki akses layak terhadap air bersih, sanitasi, serta tempat tinggal. Ketergantungan terhadap program bantuan pemerintah juga sangat minimal, menandakan bahwa daerah-daerah ini sedang berada dalam transisi menuju kesejahteraan penuh.

#### 5. Cluster 4 (Sejahtera)

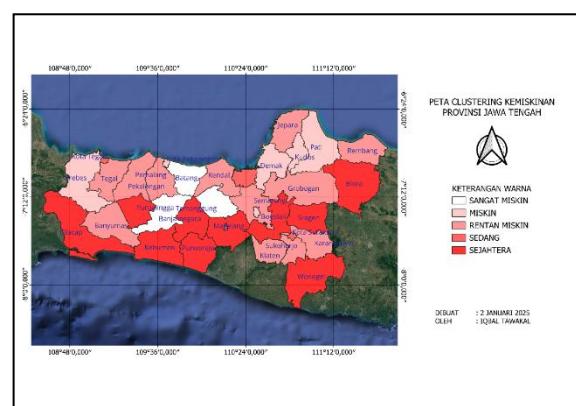
Wilayah-wilayah yang tergolong dalam klaster ini menunjukkan tingkat kesejahteraan tertinggi dibandingkan klaster lainnya. Terdiri dari Kabupaten Cilacap, Purbalingga, Kebumen, Purworejo, Wonosobo, Magelang, Boyolali, Wonogiri, Sragen, dan Blora, kelompok ini ditandai oleh hampir tidak adanya populasi miskin secara signifikan. Akses masyarakat terhadap layanan dasar sangat baik, sebagian besar rumah tangga memiliki tempat tinggal milik sendiri, dan ketergantungan terhadap bantuan sosial sangat rendah. Ciri-ciri ini menunjukkan tingkat kemandirian ekonomi yang kuat dan kualitas hidup yang lebih mapan.

Gambar 7 berikut menunjukkan visualisasi data hasil dari clustering data sebelumnya. Pada diagram, setiap titik menunjukkan area tertentu, dan warna yang berbeda menunjukkan kelompok atau cluster yang berbeda. Oleh karena itu, diagram ini memberikan gambaran pola distribusi karakteristik antar wilayah yang lebih jelas dan mudah dipahami.



Gambar 7. Scatter Plot

Peta di bawah ini menunjukkan hasil analisis clustering terhadap data kemiskinan di Provinsi Jawa Tengah. Warna-warna pada peta mewakili kelompok atau klaster yang memiliki tingkat kemiskinan yang berbeda. Visualisasi ini disajikan melalui Gambar 8.



Gambar 8. Peta Pemetaan

Peta pemetaan di atas menunjukkan bahwa ada perbedaan dalam distribusi tingkat kesejahteraan di Jawa Tengah. Peta ini juga dapat digunakan sebagai dasar untuk membuat program pembangunan yang lebih tepat sasaran dan efektif untuk mengurangi kemiskinan di Jawa Tengah, dan upaya pengentasan kemiskinan harus difokuskan pada daerah yang memiliki tingkat kemiskinan yang tinggi. Namun peta ini hanya memberikan gambaran umum tentang distribusi tingkat kemiskinan di tingkat kabupaten atau kota, data dari tingkat yang lebih rendah seperti desa atau kelurahan diperlukan untuk mendapatkan gambaran yang lebih rinci.

Untuk memperjelas penyajian, hasil penelitian disajikan secara cermat agar mudah dipahami, misalnya dapat ditunjukkan dalam bentuk tabel, kurva, grafik, gambar, foto, atau bentuk lainnya sesuai keperluan secara lengkap dan jelas. Perlu diusahakan agar saat membaca hasil penelitian dalam format tersebut, pembaca tidak perlu mencari informasi terkait dari uraian dalam pembahasan. Akhir dari bagian ini memuat keterangan tentang kelebihan dan kelemahan sistem, yang dideskripsikan secara terinci.

## Kesimpulan dan Saran

Hasil penelitian menunjukkan pola distribusi kemiskinan yang signifikan di Jawa Tengah, yang ditunjukkan oleh algoritma K-Means yang digunakan dengan RapidMiner untuk menganalisis tingkat kemiskinan. Dengan nilai DBI untuk K=5 sebesar 0,009 menunjukkan kualitas pengelompokan data yang sangat baik, maka daerah di Jawa Tengah dapat dikategorikan ke dalam lima klaster: sangat miskin, miskin, rentan miskin, sedang, dan sejahtera. Hasil ini menunjukkan bahwa tingkat kemiskinan berbeda di setiap wilayah, klaster "sangat miskin" dan "miskin" memerlukan perhatian khusus dalam program pengentasan kemiskinan. Untuk memudahkan interpretasi dan pengambilan keputusan, hasil clustering ini divisualisasikan dalam bentuk peta, grafik bar, dan plot scatter. Peta distribusi menunjukkan bahwa wilayah selatan Jawa Tengah memiliki tingkat kemiskinan yang lebih rendah dibandingkan dengan wilayah utara. Penelitian ini menawarkan solusi berbasis data untuk mendukung kebijakan pemerintah untuk mengurangi kemiskinan, khususnya melalui pembagian bantuan yang lebih tepat sasaran.

Penulis memiliki beberapa saran untuk pengembangan penelitian selanjutnya :

1. Untuk meningkatkan ketepatan analisis, penelitian berikutnya dapat menggunakan data pada tingkat desa atau kelurahan.
2. Untuk mendapatkan gambaran yang lebih baik, masukkan indikator tambahan seperti akses ke layanan kesehatan dan pendidikan.
3. Untuk membandingkan hasil dan meningkatkan validitas, gunakan metode clustering tambahan seperti fuzzy C-means.

4. Menggabungkan data dari berbagai lembaga (BPS, Dinas Sosial, dan Dinas Kesehatan) untuk memastikan kebijakan berbasis data yang lebih luas.

## Daftar Pustaka

- [1] A. Khalif *dkk.*, "Klasterisasi Tingkat Kemiskinan di Indonesia Menggunakan Algoritma K-Means."
- [2] W. Setya dan A. Nugraha, "Clustering Pemetaan Tingkat Kemiskinan di Provinsi Jawa Barat Menggunakan Algoritma K-Means," *Jurnal Ilmiah Wahana Pendidikan, Januari*, vol. 2023, no. 2, hlm. 234–244, doi: 10.5281/zenodo.7567622.
- [3] S. N. Mayasari dan J. Nugraha, "Implementasi K-Means Cluster Analysis untuk Mengelompokkan Kabupaten/Kota Berdasarkan Data Kemiskinan di Provinsi Jawa Tengah Tahun 2022," 2023. [Daring]. Tersedia pada: <https://jateng.bps.go.id/>.
- [4] D. Fuji Astri, "CLUSTERING PENDUDUK MISKIN MENGGUNAKAN ALGORITMA K-MEANS PADA WILAYAH JAWA BARAT," 2024.
- [5] Irfan dan L. Faizal, "Implementasi Algoritma K-Means untuk Mengelompokkan Data Tingkat Kemiskinan di Sulawesi Selatan Berdasarkan Kota/Kabupaten," *Jurnal Ilmiah Sistem Informasi dan Teknik Informatika (JISTI)*, vol. 7, no. 2, hlm. 261–269, Okt 2024, doi: 10.57093/jisti.v7i2.220.
- [6] R. Astuti dan F. M. Basysyar, "PENGELOMPOKKAN JUMLAH PENDUDUK MISKIN DI JAWA BARAT MENGGUNAKAN ALGORITMA K-MEANS," 2024. [Daring]. Tersedia pada: <https://data.go.id/home>.
- [7] K. Balai, S. Daerah, B. Diy Dengan Badan, P. Statistik, dan P. Diy, "LAPORAN AKHIR ANALISIS KRITERIA DAN INDIKATOR KEMISKINAN MULTIDIMENSI UNTUK DIAGNOSTIK KEMAJUAN DAERAH DI DAERAH ISTIMEWA YOGYAKARTA," 2017.
- [8] M. M. Alif, F. Ramdani, dan W. Purnomo, "Pengembangan Plugin QGIS Untuk Mengakses Peta Geologis Seluruh Indonesia," 2020. [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id>
- [9] E. Dwiguna dan A. Bahtiar, "PENERAPAN DATA MINING UNTUK MENENTUKAN PENERIMA BANTUAN BLT MENGGUNAKAN METODE CLUSTERING K-MEANS PADA DESA PAMULIHAN," 2024.
- [10] I. Manfaati Nur, M. Rizky, dan S. Putri Milasari, "Pengelompokan Tingkat Kemiskinan di Provinsi Jawa Barat dengan Metode K-Means Clustering Info Artikel,"

vol. 1, no. 2, hlm. 51–61, 2023, doi: 10.26714/jodi.

- [11] N. Sukarno Wijaya, M. Jajuli, dan B. A. Dermawan, “PENERAPAN ALGORITMA K-MEANS CLUSTERING DALAM MENENTUKAN DAERAH PRIORITAS PENANGANAN KEMISKINAN DI WILAYAH JAWA TIMUR,” 2024.
- [12] U. Saidata Aesyi, “ANALISIS TINGKAT KEBERMANFAATAN MYPERTAMINA MENGGUNAKAN K-MEANS CLUSTERING,” 2023.
- [13] M. Astriani, M. Hamu, dan A. C. Talakua, “Optimalisasi Manajemen Bantuan Sosial Dengan Implementasi Data Mining Menggunakan Algoritma K-Means Clustering,” 2024.