

KLASTERISASI KARAKTERISTIK UMKM MENGGUNAKAN ALGORITMA K-MEANS DI WILAYAH KABUPATEN KARAWANG

Ni Made Ayu Gayatri Pratiwi, Shofa Shofiah Hilabi, Elfina Novalia, Tukino

Sistem Informasi, Universitas Buana Perjuangan Karawang

Jalan HS. Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Kabupaten Karawang, Jawa Barat, Indonesia

si22.nipratiwi@mhs.ubpkarawang.ac.id

ABSTRAK

Pertumbuhan ekonomi lokal dipengaruhi oleh usaha mikro, kecil, dan menengah (UMKM), namun persebarannya di Kabupaten Karawang masih belum merata. Penelitian ini bertujuan untuk mengelompokkan UMKM guna mengidentifikasi perbedaan karakteristik antarwilayah. Data yang digunakan berupa jumlah UMKM pada sektor industri kecil, industri rumah tangga, jasa, dan perdagangan periode 2019–2025. Metode yang digunakan adalah K-Means Clustering dengan penentuan jumlah kluster menggunakan metode Elbow serta evaluasi menggunakan Silhouette Score. Hasil penelitian menunjukkan bahwa jumlah kluster optimal adalah empat dengan nilai Silhouette Score sebesar 0,436 yang menunjukkan kualitas pengelompokan cukup baik. Setiap kluster memiliki karakteristik berbeda, yaitu aktivitas UMKM tinggi, menengah, dan rendah dengan dominasi sektor tertentu. Selain itu, satu kecamatan dapat berada pada kluster yang berbeda pada tahun yang berbeda, yang mencerminkan dinamika perkembangan UMKM. Hasil visualisasi juga menunjukkan bahwa persebaran UMKM di Kabupaten Karawang masih belum merata. Penelitian ini diharapkan dapat menjadi dasar dalam penyusunan kebijakan pengembangan UMKM yang lebih terarah.

Kata kunci : *UMKM, Data Mining, K-Means Clustering, Klasterisasi Wilayah, Power BI*

1. PENDAHULUAN

Usaha mikro, kecil, dan menengah (UMKM) memegang peranan penting dalam perekonomian nasional karena dapat meningkatkan lapangan kerja dan pertumbuhan ekonomi daerah. Keberadaan sektor UMKM juga berkontribusi terhadap ketahanan ekonomi, terutama mengingat kondisi ekonomi yang tidak stabil [1]. Kabupaten Karawang, sebagai salah satu wilayah industri di Provinsi Jawa Barat, mengalami peningkatan jumlah UMKM yang cukup signifikan setiap tahunnya. Namun, pertumbuhan tersebut belum merata di seluruh kecamatan sehingga menimbulkan perbedaan karakteristik dan potensi ekonomi antarwilayah.

Selama ini, analisis kondisi UMKM pada tingkat wilayah umumnya masih dilakukan secara deskriptif berdasarkan jumlah unit usaha atau kategori usaha yang ada, sehingga belum mampu menggambarkan pola kesamaan karakteristik antarwilayah secara menyeluruh. Kondisi ini menunjukkan perlunya pendekatan berbasis data yang lebih objektif untuk mengidentifikasi pola tersebut. Dalam hal ini, teknik data mining dapat dimanfaatkan untuk mengolah data dalam jumlah besar menjadi informasi yang lebih bermakna dalam mendukung pengambilan keputusan [2]. Salah satu metode yang banyak digunakan adalah algoritma K-Means, yang bekerja dengan mengelompokkan data berdasarkan tingkat kemiripan terhadap pusat kluster sehingga mampu mengidentifikasi pola tertentu dalam data [3][4]. Selain itu, perkembangan UMKM yang terjadi dari waktu ke waktu menunjukkan adanya perubahan pola aktivitas ekonomi pada setiap wilayah [5].

Berdasarkan hal tersebut, penelitian ini difokuskan pada klasterisasi UMKM di Kabupaten

Karawang untuk melihat perbedaan karakteristik antarwilayah yang selama ini belum tergambarkan secara merata. Variabel yang digunakan meliputi jumlah pelaku usaha pada industri rumah tangga, perdagangan, dan jasa. Temuan dalam penelitian ini diharapkan dapat menemukan potensi ekonomi wilayah dan memberikan informasi yang dapat digunakan pemerintah daerah dalam merancang kebijakan pengembangan UMKM secara lebih terarah.

2. TINJAUAN PUSTAKA

2.1. Usaha Mikro, Kecil, dan Menengah

Perusahaan-perusahaan yang menguntungkan dan beroperasi di berbagai sektor industri dikenal sebagai Usaha Mikro, Kecil, dan Menengah (UMKM). UMKM memiliki potensi untuk secara signifikan meningkatkan lapangan kerja dan berkontribusi pada pertumbuhan perekonomian nasional [1]. Selain itu, sektor UMKM juga dinilai cukup tahan terhadap perubahan kondisi ekonomi sehingga berperan penting dalam menjaga stabilitas perekonomian daerah [2]. Peningkatan jumlah UMKM yang terjadi di berbagai wilayah menunjukkan tingginya partisipasi masyarakat dalam kegiatan ekonomi produktif, namun distribusi pertumbuhannya tidak selalu merata sehingga diperlukan analisis untuk memahami karakteristiknya secara lebih mendalam [3].

2.2. Data Mining

Data Mining dapat menemukan pola-pola tersembunyi dalam kumpulan data yang sangat besar merupakan tujuan dari penambangan data, sehingga pola-pola tersebut dapat menjadi landasan dalam pengambilan keputusan [5]. Teknik ini memungkinkan data mentah diubah menjadi data yang

lebih relevan. sehingga membantu berbagai bidang, termasuk ekonomi dan bisnis, dalam memahami karakteristik suatu objek secara lebih objektif [6]. Pemanfaatan *data mining* menjadi penting karena mampu mengidentifikasi pola tersembunyi yang sulit ditemukan melalui analisis deskriptif biasa.

Secara umum, proses penambangan data terdiri dari beberapa langkah, salah satunya adalah pemilihan data, membersihkan data, dan mengubahnya, proses analisis menggunakan *algoritma* tertentu, serta evaluasi hasil yang diperoleh [7]. Algoritma pengelompokan digunakan dalam penambangan data untuk membagi data menjadi beberapa kelompok berdasarkan karakteristik yang sama [8]. Pada konteks UMKM, metode ini dapat membantu mengelompokkan wilayah usaha sehingga potensi ekonomi tiap daerah dapat dianalisis dengan lebih jelas[9].

2.3. Algoritma K-Means

Ketika teknik K-Means digunakan untuk pengelompokan, setiap titik data akan dikelompokkan ke dalam kluster yang paling dekat dengan pusat. Karena metode komputasinya yang sederhana, kemudahan implementasinya, dan kemampuannya untuk menghasilkan pengelompokan data yang efisien di berbagai bidang studi, teknik K-Means banyak digunakan [10].

Sampai konvergensi tercapai, algoritma K-Means secara berulang menentukan jumlah kluster yang ada dan memilih pusat kluster pertama, menghitung jarak antara pusat kluster dan setiap titik data, lalu memperbarui nilai pusat kluster berdasarkan rata-rata data dari kluster ini [11] Perhitungan jarak umumnya menggunakan jarak geometri, yang dapat dirumuskan sebagai berikut:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^n (x_{ik} - c_{jk})^2} \quad (1)$$

Keterangan:

- $d(x_i, c_j)$ = jarak antara centroid ke-j dan data ke-i
- x_{ik} = nilai atribut k data ke-i
- c_{jk} = nilai atribut centroid ke-j ke-k
- n = jumlah atribut

Untuk mengidentifikasi kesamaan dalam potensi ekonomi di berbagai wilayah, teknik K-Means digunakan dalam penelitian ini guna mengelompokkan kecamatan-kecamatan di Kabupaten Karawang berdasarkan jumlah pelaku UMKM.

2.4. Bahasa Pemrograman Python

Python termasuk dalam kategori bahasa pemrograman yang paling populer untuk pengolahan data yang populer karena menyediakan beragam pustaka pendukung dan mudah digunakan dalam implementasi analisis data[13]. Beberapa pustaka antara lain *NumPy*, yang biasanya digunakan dalam proses clustering, *Pandas*, dan *Scikit-learn* yang

menyediakan fungsi perhitungan *algoritma* secara efisien serta mendukung proses analisis data secara terstruktur [14]. Selain itu, *Python* juga mampu menghasilkan visualisasi data yang informatif sehingga mempermudah interpretasi hasil analisis. Kemudahan penggunaan serta fleksibilitas yang dimiliki menjadikan *Python* sebagai salah satu alat terbaik untuk menggunakan algoritma K-Means pada penelitian ini.

2.5. Power BI

Power BI adalah software visualisasi data yang digunakan untuk menyajikan informasi dalam bentuk *dashboard* interaktif sehingga memudahkan pengguna dalam memahami pola data. Aplikasi ini mampu mengolah data numerik menjadi tampilan visual seperti grafik, tabel, maupun peta sehingga informasi yang dihasilkan lebih mudah dianalisis dan diinterpretasikan [15]. Penelitian ini menggunakan Power BI sebagai alat visualisasi untuk menampilkan hasil klasterisasi wilayah berdasarkan karakteristik UMKM sehingga perbedaan antar kluster dapat terlihat dengan lebih jelas serta mendukung proses pengambilan keputusan berbasis data [16].

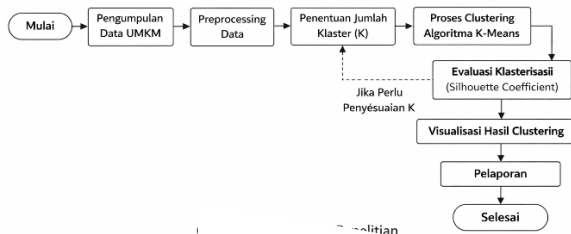
2.6. Penelitian Terkait

Beberapa penelitian terdahulu menunjukkan bahwa algoritma K-Means efektif digunakan dalam analisis persebaran UMKM. Penelitian oleh Wafi dan Arianto [9] menghasilkan tiga kluster UMKM dengan kualitas pengelompokan yang baik berdasarkan nilai Silhouette Coefficient sebesar 0,73 dan indeks Davies-Bouldin sebesar 0,29. Penelitian lain oleh Arrosyad et al. [4] juga menghasilkan pembagian kluster UMKM ke dalam kategori rendah, sedang, dan tinggi dengan evaluasi menggunakan indeks Davies-Bouldin. Selanjutnya, Sudrajat et al. [17] memperoleh lima kluster optimal dengan tingkat akurasi sebesar 66,7% dari total data yang digunakan. Selain itu, Wikarta dan Defiyanti [2] mengelompokkan UMKM menjadi tiga kluster dengan nilai Silhouette sebesar 0,58 yang menunjukkan kualitas pengelompokan yang cukup baik. Penelitian oleh Ameliana et al. [11] bahkan menghasilkan sepuluh kluster dengan nilai Davies-Bouldin Index sebesar 0,527.

Berdasarkan penelitian terdahulu tersebut, dalam hal mengelompokkan data berdasarkan atribut tertentu, algoritma K-Means bekerja dengan efektif. Namun, sebagian penelitian masih menggunakan cakupan wilayah yang luas seperti tingkat provinsi atau kota sehingga belum menggambarkan kondisi secara rinci pada tingkat kecamatan. Akibatnya, penelitian ini dilakukan dengan mengutamakan pada klasterisasi wilayah kecamatan di Kabupaten Karawang menggunakan data terbaru sehingga diharapkan mampu memberikan gambaran yang lebih khusus tentang ciri-ciri UMKM yang berbeda pada setiap wilayah.

3. METODE PENELITIAN

Dalam penelitian ini, kecamatan-kecamatan di Kabupaten Karawang dikelompokkan berdasarkan karakteristik UMKM dengan menggunakan teknik K-Means. Proses analisis dimulai dari pengumpulan data, dilanjutkan dengan *pre-processing*, penerapan metode *clustering*, hingga evaluasi terhadap hasil pengelompokan. Data yang digunakan merupakan data jumlah pelaku UMKM berdasarkan kategori usaha yang diperoleh dari sumber data resmi pemerintah.



Gambar 1. Alur penelitian

Langkah-langkah penelitian yang dilakukan dalam pengelompokan regional berdasarkan karakteristik UMKM dengan menggunakan teknik K-Means digambarkan pada Gambar 1. Langkah-langkah yang digunakan dalam penelitian ini adalah sebagai berikut:

- Pengumpulan Data**
Tahapan ini berfokus pada pengumpulan data mengenai jumlah pelaku UMKM menurut kategori usaha yang diperoleh dari sumber resmi pemerintah.
- Preprocessing Data**
Data yang telah dikumpulkan selanjutnya diproses melalui pembersihan serta penyesuaian bentuk agar sesuai untuk kebutuhan analisis.
- Penentuan Jumlah Kluster (K)**
Tahapan ini berfokus pada penetapan jumlah kluster ini digunakan dalam metode K-Means untuk memperoleh hasil pengelompokan yang optimal.
- Algoritma Pengelompokan K-Means**
Dengan menerapkan metode K-Means, informasi yang telah diolah digunakan untuk mengelompokkan area berdasarkan karakteristik yang sama.
- Evaluasi Klasterisasi**
Pendekatan Koefisien Siluet digunakan sebagai indikator evaluasi kluster untuk menilai kualitas hasil pengelompokan.
- Visualisasi Hasil Clustering**
Hasil *clustering* disajikan dalam bentuk grafik atau *dashboard* untuk mempermudah proses interpretasi.
- Pelaporan**
Tahap akhir penelitian adalah penyusunan laporan tergantung pada hasil analisis yang telah dilakukan.

Secara umum, tahapan-tahapan penelitian telah dirancang secara sistematis untuk memastikan bahwa seluruh prosedur mulai dari analisis kluster hingga pemrosesan data dapat diselesaikan dengan efektif. Setiap tahap proses, mulai dari pengolahan data awal hingga visualisasi hasil, saling terintegrasi untuk memberikan gambaran yang lebih jelas dan akurat mengenai kondisi UMKM di Kabupaten Karawang.

4. HASIL DAN PEMBAHASAN

4.1. Deskripsi Data

Usaha Mikro, Kecil, dan Menengah (UMKM) yang menjadi objek data dalam penelitian ini yang berada pada wilayah kecamatan di Kabupaten Karawang dengan rentang tahun 2019 sampai 2025. Data tersebut berisi informasi mengenai jumlah unit usaha berdasarkan beberapa kategori sektor usaha pada masing-masing kecamatan. Data ini digunakan sebagai dasar dalam melakukan analisis pengelompokan wilayah berdasarkan karakteristik UMKM yang dimiliki.

Penggunaan data dalam beberapa tahun dilakukan agar analisis yang dihasilkan tidak hanya menggambarkan kondisi pada satu waktu tertentu, tetapi juga dapat merepresentasikan perkembangan UMKM secara lebih menyeluruh. Dengan demikian, pola yang terbentuk dari hasil *clustering* diharapkan lebih mencerminkan kondisi asli di lapangan.

Setiap kecamatan memiliki komposisi jumlah UMKM yang berbeda pada masing-masing sektor. Perbedaan ini menunjukkan bahwa aktivitas ekonomi di setiap wilayah tidak konsisten, sehingga diperlukan metode analitis yang dapat mengelompokkan kecamatan berdasarkan ciri-ciri yang sama.

4.2. Preprocessing

Data tersebut melalui tahap prapemrosesan sebelum masuk ke prosedur pengelompokan. Tahap ini bertujuan untuk meningkatkan kesesuaian data guna digunakan dalam proses analisis serta mengurangi kemungkinan terjadinya kesalahan perhitungan algoritmik.

Langkah awal adalah melakukan penyesuaian struktur data agar setiap kecamatan memiliki atribut yang sama untuk setiap kategori sektor UMKM. Data yang awalnya masih terpisah berdasarkan kategori usaha kemudian diolah menggunakan teknik *pivot* sehingga menghasilkan satu *dataset* yang memuat jumlah industri kecil, industri rumah tangga, jasa, dan perdagangan pada setiap kecamatan. Hasil transformasi data dapat dilihat pada Gambar 2.

tahun	nama_kecamatan	INDUSTRI KECIL	INDUSTRI RUMAH TANGGA	JASA	PERDAGANGAN
2019	BANYUSARI	861	587	628	3148
2019	BATUJAYA	849	713	901	6749
2019	CIAMPEL	490	893	581	3154
2019	CIBUAYA	478	366	607	1160
2019	CIKAMPEK	2129	1034	826	10520
2019	CILAMAYA KULON	1052	789	1144	6562

Gambar 2. Hasil transformasi data UMKM menggunakan teknik *pivot*

Berdasarkan Gambar 2, terlihat bahwa data yang sebelumnya berbentuk tidak terstruktur telah diubah menjadi *format* tabular, dimana setiap baris merepresentasikan kecamatan pada tahun tertentu dan setiap kolom menunjukkan kategori sektor UMKM. Selain itu, dilakukan perhitungan proporsi pada setiap sektor UMKM terhadap total UMKM untuk menggambarkan kontribusi masing-masing sektor pada setiap kecamatan.

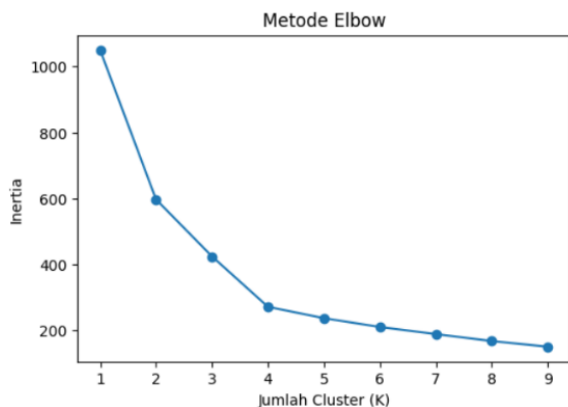
Selanjutnya, dilakukan transformasi data menggunakan fungsi *logaritma*. Transformasi ini bertujuan untuk mengurangi perbedaan nilai yang terlalu besar antar kecamatan. Pada data jumlah UMKM, terdapat beberapa wilayah dengan nilai yang jauh lebih tinggi dibandingkan wilayah lainnya, sehingga transformasi *log* digunakan untuk membantu menstabilkan distribusi data.

Setelah proses transformasi dilakukan, data menjadi lebih seimbang sehingga perhitungan jarak pada *algoritma K-Means* tidak didominasi oleh variabel dengan nilai yang besar. Tahapan *pre-processing* ini bertujuan untuk meningkatkan kualitas hasil *clustering* sehingga kelompok yang terbentuk dapat merepresentasikan kondisi UMKM secara lebih akurat.

4.3. Penentuan Jumlah Kluster dan Evaluasi Model

Karena hal ini memengaruhi baik hasil pengelompokan maupun interpretasi setiap kelompok yang dihasilkan, menentukan jumlah kelompok merupakan langkah penting dalam penerapan *algoritma K-Means*. Oleh karena itu, jumlah kelompok yang paling sesuai dengan karakteristik data harus ditentukan melalui prosedur evaluasi.

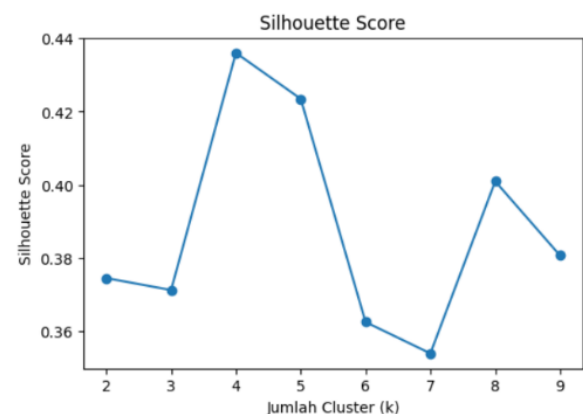
Pendekatan *Elbow* digunakan dalam penelitian ini untuk menentukan jumlah kluster. Teknik ini digunakan untuk melacak variasi nilai inersia sehubungan dengan jumlah kluster yang diuji. Untuk menentukan titik di mana penurunan nilai inersia mulai melambat, pengujian dilakukan dengan bereksperimen menggunakan berbagai nilai K. Hasilnya kemudian ditampilkan secara grafis.



Gambar 3. Grafik hasil kluster menggunakan metode *elbow*

Berdasarkan hasil pengujian, terlihat bahwa penurunan inersia mulai tidak signifikan setelah jumlah kluster mencapai empat. Hal ini menunjukkan bahwa kualitas pengelompokan tidak meningkat secara signifikan jika jumlah kelompok ditingkatkan melebihi batas tersebut.

Penilaian dengan menggunakan Skor Silhouette dilakukan untuk memvalidasi temuan-temuan ini. Teknik Hal ini digunakan untuk mengukur seberapa dekat data-data dalam suatu kluster dan seberapa jauh jarak antar kluster. Kualitas kluster yang lebih baik ditandai dengan nilai Skor Silhouette yang mendekati 1, yang berkisar antara -1 hingga 1.



Gambar 4. Grafik hubungan jumlah kluster (K) dengan nilai *Silhouette Score*

Sebagaimana ditunjukkan pada Gambar 4, jelas bahwa nilai *Silhouette Score* mengalami peningkatan hingga mencapai nilai tertinggi pada jumlah kluster sebanyak 4. Setelah titik tersebut, Skor Siluet menurun pada jumlah kluster berikutnya. Hal ini menunjukkan bahwa penggunaan empat kluster menghasilkan tingkat pemisahan kelompok yang lebih tinggi berkaitan dengan jumlah kelompok lainnya, sehingga mampu merepresentasikan struktur data dengan lebih optimal.

Tabel 2. Perbandingan nilai inersia dan silhouette

Jumlah Kluster	Inertia	Silhouette
2	596.786	0.374
3	424.154	0.371
4	271.368	0.436

Berdasarkan Tabel 2, nilai inersia mengalami penurunan seiring bertambahnya jumlah kluster, namun penurunan mulai melambat pada K=4 yang mengindikasikan titik optimal berdasarkan metode *Elbow*. Namun, nilai *Silhouette Score* maksimum ($\pm 0,436$) juga dicapai pada K=4. Hal ini menunjukkan bahwa kedua pendekatan tersebut menghasilkan hasil yang konsisten; oleh karena itu, dalam penelitian ini dipilih empat kluster.

4.4. Distribusi Kecamatan pada Setiap Klaster

Hasil *clustering* menunjukkan bahwa jumlah kecamatan pada setiap klaster tidak merata. Beberapa klaster memiliki jumlah anggota yang lebih banyak, sementara klaster lainnya terdiri dari jumlah kecamatan yang lebih sedikit.

Tabel 3. Hasil distribusi kecamatan

Klaster	Jumlah Kecamatan
Klaster 0	60
Klaster 1	46
Klaster 2	40
Klaster 3	64

Berdasarkan hasil *clustering* yang ditunjukkan pada Tabel 3, diketahui bahwa jumlah data pada setiap klaster memiliki perbedaan. Klaster 3 memiliki jumlah data terbanyak yaitu sebanyak 64 data, diikuti oleh klaster 0 sebanyak 60 data. Sementara itu, klaster 1 dan klaster 2 masing-masing memiliki jumlah data sebanyak 46 dan 40 data.

Perlu diperhatikan bahwa informasi yang digunakan dalam penelitian ini berasal dari rangkaian data waktu dari tahun 2019 hingga 2025. Oleh karena itu, satu kecamatan dapat muncul lebih dari satu kali dalam dataset karena direpresentasikan pada tahun yang berbeda. Kondisi ini memungkinkan satu kecamatan berada pada klaster yang berbeda di setiap tahunnya, tergantung pada perubahan jumlah dan karakteristik UMKM yang dimiliki.

Perbedaan dalam jumlah anggota di antara setiap klaster menunjukkan bahwa fitur tertentu lebih dominan dibandingkan karakteristik lainnya. Klaster dengan jumlah anggota yang lebih banyak mengindikasikan bahwa pola tersebut lebih umum ditemukan pada data, sedangkan klaster dengan jumlah anggota yang lebih sedikit menunjukkan karakteristik yang lebih spesifik.

4.5. Karakteristik Setiap Klaster

Nilai rata-rata untuk setiap sektor UMKM dalam setiap klaster dianalisis guna mengidentifikasi karakteristik masing-masing klaster. Tabel 4 menyajikan hasil perhitungan rata-rata untuk variabel-variabel dalam setiap klaster.

Tabel 4. Rata-rata karakteristik UMKM di setiap klaster

Klaster	Industri Kecil	Industri RT	Jasa	Perdagangan
0	1697	1062	1193	10216
1	91	440	504	963
2	1247	711	1232	3592
3	229	724	679	3238

Berdasarkan Tabel 4, klaster 0 memiliki nilai rata-rata tertinggi pada seluruh sektor, terutama pada sektor perdagangan yang mencapai nilai paling

dominan. Hal ini menunjukkan bahwa klaster 0 merupakan wilayah dengan aktivitas UMKM yang sangat tinggi dan dapat dikategorikan sebagai pusat kegiatan ekonomi.

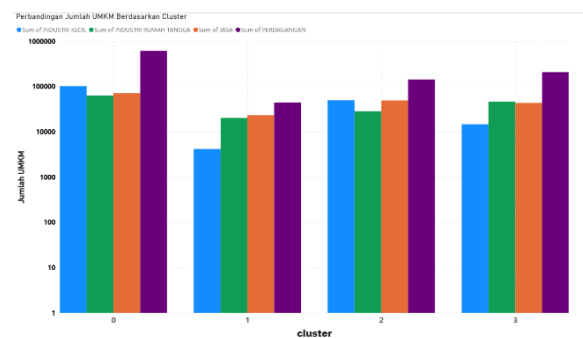
Klaster 2 menunjukkan nilai yang cukup tinggi pada sektor industri kecil dan jasa, namun tidak setinggi klaster 0. Hal ini menunjukkan bahwa sektor UMKM, yang mencakup berbagai sektor ekonomi, sedang berkembang pesat di wilayah-wilayah yang tergabung dalam klaster ini.

Sementara itu, klaster 3 memiliki nilai yang berada pada tingkat menengah pada sebagian besar sektor. Kondisi ini menunjukkan bahwa meskipun kegiatan UMKM belum terlalu meluas, namun cukup merata di seluruh wilayah klaster tersebut.

Klaster 1 memiliki nilai rata-rata paling rendah pada hampir seluruh sektor, sehingga dapat dikategorikan sebagai daerah dengan aktivitas UMKM relatif rendah dan masih memiliki potensi untuk berkembang.

4.6. Visualisasi

Setelah proses *clustering* dilakukan, tahap selanjutnya adalah menunjukkan hasil analisis melalui visualisasi. Visualisasi dilakukan menggunakan *Microsoft Power BI* untuk mempermudah memahami pola distribusi UMKM serta perbedaan karakteristik antarklaster pada setiap kecamatan di Kabupaten Karawang. Penyajian dalam bentuk *dashboard* interaktif ini memungkinkan data yang kompleks dapat ditampilkan secara lebih informatif dan mudah dipahami.



Gambar 5. Visualisasi jumlah UMKM pada setiap kecamatan

Berdasarkan Gambar 5, terlihat bahwa terdapat perbedaan jumlah UMKM yang cukup jelas pada setiap sektor antarklaster. Klaster 0 menunjukkan nilai tertinggi, menunjukkan bahwa di hampir setiap sektor industri, bidang-bidang dalam kategori ini memiliki tingkat aktivitas ekonomi tertinggi, terutama sektor perdagangan. Sementara itu, klaster 1 memiliki jumlah UMKM yang relatif paling rendah pada hampir seluruh sektor. Klaster 2 dan klaster 3 berada pada tingkat menengah dengan karakteristik yang lebih seimbang antar sektor usaha.

[illegible]

Gambar 6. Peta persebaran klaster UMKM di kabupaten Karawang

Dengan adanya visualisasi berbasis *dashboard* ini, hasil *clustering* tidak hanya dapat dipahami secara numerik, tetapi juga secara visual, sehingga dapat memfasilitasi pengambilan keputusan yang lebih baik terkait pertumbuhan UMKM di Kabupaten Karawang serta membantu mengidentifikasi pola distribusi UMKM.

Berdasarkan hasil penelitian, pendekatan K-Means berhasil mengelompokkan data UMKM Kabupaten Karawang untuk periode tahun 2019–2025 menjadi empat kluster, dengan skor siluet sebesar $\pm 0,436$, yang menunjukkan kualitas pengelompokan yang cukup baik. Jumlah kluster ditentukan dengan menggunakan pendekatan Elbow dan skor siluet.

DAFTAR PUSTAKA

- 4646

- Colab Untuk Keterampilan Teknologi Siswa Smk,” Vol. 9, No. 1, Pp. 127–134, 2025, Doi: 10.37729/Abdimas.V9i1.5705.
- [8] E. Novalia And A. Voutama, “Prediction Of Rice Field Planted Area With Crisp-Dm Using Classification And Regression Tree (Cart) Algorithms,” 2023.
- [9] F. M. Wafi And D. B. Arianto, “Implementasi Metode K-Means Clustering Dalam Analisis Persebaran Umkm Di Jawa Barat,” 2022.
- [10] N. Syifa And R. N. Fahmi, “Implementasi Metode K-Means Clustering Dalam Analisis Persebaran Umkm Di Jawa Barat,” *Joins (Journal Of Information System)*, Vol. 6, No. 2, Pp. 211–220, Dec. 2021, Doi: 10.33633/Joins.V6i2.5310.
- [11] N. Ameliana, N. Suarna, And W. Prihartono, “Analisis Data Mining Pengelompokkan UMKM Menggunakan Algoritma K-Means Clustering Di Provinsi Jawa Barat,” 2024.
- [12] T. M. Ghazal *Et Al.*, “Performances Of K-Means Clustering Algorithm With Different Distance Metrics,” *Intelligent Automation And Soft Computing*, Vol. 30, No. 2, Pp. 735–742, 2021, Doi: 10.32604/Iasc.2021.019067.
- [13] D. Aryani, B. Irawan, dan A. Bahtiar, “Implementasi Data Mining Pada Data Penjualan Pakaian Menggunakan Algoritma K-Means Dengan Optimize Parameter Grid,” *Journal Mahasiswa Teknik Informatika*, Vol. 8, No. 2, 2024.
- [14] D. A. Manalu And G. Gunadi, “Implementasi Metode Data Mining K-Means Clustering Terhadap Data Pembayaran Transaksi Menggunakan Bahasa Pemrograman Python Pada Cv Digital Dimensi,” *Infotech: Journal Of Technology Information*, Vol. 8, No. 1, Pp. 43–54, Jun. 2022, Doi: 10.37365/Jti.V8i1.131.
- [15] H. Stiawan *Et Al.*, “Model Visualisasi Informasi Dashboard Pada Pemetaan Tanaman Obat Dan Langka Kabupaten Kediri Menggunakan Microsoft Power Bi,” 2022.
- [16] M. K. Najib dan E. M. Stefany, “Visualisasi Data Penjualan Supermarket Dengan Microsoft Power Bi Untuk Menghasilkan Insight Dan Rekomendasi,” *Jurnal Sistem Informasi (Teknofile)*, Vol. 2, Pp. 921–928, 2024.
- [17] W. Sudrajat, I. Cholid, dan J. Petrus, “Penerapan Algoritma K-Means Untuk Pengelompokan Umkm Menggunakan Rapidminer,” 2022.
- [18] Hendrik dan T. M. S. Mulyana, “Pemetaan Daerah Berdasarkan Jenis Usaha Umkm Dengan Algoritma K-Means Di Jawa Barat (Regional Mapping Based On The Type Of Umkm Business With K-Means Algorithm In West Java),” 2023. [Online]. Available: [Http://jtika.If.Unram.Ac.Id/Index.Php/Jtika/](http://jtika.if.unram.ac.id/index.php/jtika/)