



Pengelompokkan Tingkat Pengangguran Terbuka di Jawa Tengah Menggunakan Metode K – Means Clustering

Fatchi Arifatul Ulya¹, Ammil Nur Abdullah^{2*}, Tio Aisyah Hanan³, Indah Manfaati Nur⁴

¹Universitas Muhammadiyah Semarang, Indonesia

²Universitas Muhammadiyah Semarang, Indonesia

³Universitas Muhammadiyah Semarang, Indonesia

⁴Universitas Muhammadiyah Semarang, Indonesia

DOI: <https://doi.org/10.26714/jodi>

Info Artikel

Sejarah Artikel:

Disubmit 14 Juli 2023

Direvisi 20 Desember 2023

Disetujui 22 Desember 2023

Keywords:

Data Mining; K-Means

Clustering; Open Unemployment.

Abstrak

Pengangguran merupakan sebuah individu yang tidak memiliki aktivitas serta tidak menghasilkan pendapatan. Tujuan dari penelitian ini mengelompokkan tingkat pengangguran terbuka menurut Provinsi Jawa Tengah. Jenis penelitian yang dilakukan menggunakan metode kuantitatif. Metode kuantitatif merupakan penelusuran data dengan cara sekunder atau dapat dikatakan memanfaatkan data yang bersumber dari Badan Pusat Statistik Jawa Tengah dengan memanfaatkan 35 Kabupaten/Kota. Dalam kasus ini dalam analisis datanya menggunakan data mining dengan pengelompokan metode k-means mengambil 3 indikator, yaitu Jumlah Penduduk, UMK, dan Jumlah Industri guna menganalisis kasus tersebut. Dalam penentuan jumlah *cluster* optimal dengan memanfaatkan metode Elbow menggunakan grafik Elbow. Grafik Elbow menghasilkan titik yang menurun drastis berbentuk siku pada jumlah *cluster* $k = 4$. Analisis *cluster* hanya sampai 2 iterasi menghasilkan pusat *cluster*, jarak dan *cluster* yang sama serta tidak terjadi perubahan. Hasil pengelompokan pada *cluster* 1 sebanyak 9 Kabupaten/Kota, *cluster* 2 sebanyak 8 Kabupaten/Kota, *cluster* 3 sebanyak 12 Kabupaten/Kota, dan *Cluster* 4 sebanyak 6 Kabupaten/Kota Jawa Tengah. Nilai akurasi dari *cluster* yang terbentuk sebesar 66.8%.

Abstract

Unemployed are individuals who have no activity and do not generate income. The purpose of this research is to classify the level of open response according to Central Java Province. This type of research was carried out utilizing quantitative methodologies. The quantitative method is data tracking in a secondary way or it can be said that it utilizes data sourced from the Central Bureau of Statistics of Central Java by utilizing 35 Regencies/ Cities. In this case, the data analysis applies k-means data mining technique grouping method, taking 3 indicators, namely Population, UMK, and Number of Industries to analyze the case. In increasing the optimal number of clusters by utilizing the Elbow method using the Elbow graph. The Elbow graph produces points that decrease drastically in the shape of an elbow at the number of clusters $k=4$. Cluster analysis only up to 2 iterations produces the same cluster center, distance and cluster and no changes occur. The results of grouping in cluster 1 were 9 regencies / cities, cluster 2 was 8 regencies / cities, cluster 3 was 12 regencies / cities, and cluster 4 was 6 regencies / cities in Central Java. The accuracy value of the cluster formed is 66.8%.

✉ Alamat Korespondensi:

E-mail: ammilnurabdullah@gmail.com

PENDAHULUAN

Sejak tahun 1980-an pengangguran terus mengalami peningkatan, yang mana sebagian besar disebabkan oleh dampak jangka pendek maupun jangka panjang [1]. Pengangguran yaitu seseorang yang sedang tidak melakukan aktivitas yang menghasilkan pendapatan, termasuk mereka yang belum memiliki pekerjaan tetapi sedang mencari pekerjaan, dan yang sedang bekerja namun tidak menghasilkan produk yang bernilai [2]. Pengangguran menjadi ancaman masalah yang ditakuti dalam sebuah negara, ketika pengangguran jumlah meningkat, sumber daya manusia yang akan terbuang sia-sia, akan mengarah pada penurunan tingkat pendapatan yang dihasilkan oleh masyarakat. Pengangguran yang ada di Indonesia dapat berasal dari beberapa faktor yang berbeda, antara lain tingginya tingkat pertumbuhan penduduk, tidak seimbangnya lapangan kerja yang tersedia, penduduk usia produktif yang tidak siap dalam menghadapi perkembangan teknologi yang pesat, tingkat pendidikan, perubahan tingkat perekonomian penduduk, dan pendapatan upah minimum di daerah tertentu [3]. Lapangan pekerjaan yang sedikit untuk pelamar tenaga kerja akan mengakibatkan timbulnya isu-isu di suatu wilayah seperti penurunan tingkat kemakmuran, efisien kerja, dan penghasilan penduduk, yang pada gilirannya akan berpengaruh terhadap kemiskinan, tingkat kejahatan, dan disparitas standar hidup [4].

Banyak sekali dampak yang ditimbulkan dari pengangguran. Tinggi rendahnya tingkat pengangguran, memperhatikan dampaknya terhadap tingkat kemiskinan disuatu wilayah [5]. Tingkat pengangguran para penduduk di Indonesia termasuk dalam kategori tinggi, meskipun mengalami penurunan dari tahun 2020 hingga 2021 [6]. Pada agustus tahun 2020, tingkat pengangguran di Indonesia sebesar 7.07%. Sedangkan pada agustus tahun 2021, tingkat pengangguran di Indonesia turun 0.58% terpantau sebesar 6.49%. Tingkat pengangguran di Provinsi Jawa Tengah berdasarkan hasil Sakernas pada februari tahun 2021, sebesar 5.96%. Hal ini menunjukkan tingkat pengangguran di Jawa tengah mengalami kenaikan sebesar 1.76% dari februari 2020. Namun mengalami penurunan pada agustus 2020 sebesar 0.52%. Tingginya pengangguran termasuk ke dalam masalah sosial dan ekonomi. Meningkatnya jumlah pengangguran yang dialami oleh penduduk yang berada di Indonesia merupakan suatu masalah yang menghambat pada pertumbuhan ekonomi di jawa tengah [7]. Pertumbuhan ekonomi di Indonesia pada tahun 2021 secara kumulatif bertumbuh dengan baik karena mencapai 3,69% lebih positif jika dibandingkan dengan kontraksi sebesar 2,07% yang terjadi pada tahun 2020 (Badan Pusat Statistik, 2021). Hal ini menunjukkan bahwa ekonomi di Indonesia pada tahun 2021 secara spasial di dominasi oleh Jawa Tengah yang memberikan kontribusi disektor ekonomi sebesar 57,89% dan kinerja disektor ekonomi dapat mengalami kenaikan sebesar 3,66%.

Tingkat Pengangguran di Jawa Tengah menggunakan teknik *clustering* yang tepat untuk digunakan yaitu Algoritma K-Means. Metode clustering atau pengelompokkan data memiliki tujuan untuk membagi dataset menjadi beberapa kelompok berdasarkan kesamaan variabel atau atribut. Dimana data dikelompokkan berdasarkan pusat *cluster (centroid)* yang paling dekat dengan data tersebut [8]. Sedangkan pada metode K-Means bertujuan untuk menstrukturkan data dengan meningkatkan kesamaan data dalam kelompok yang ditentukan dan mengurangi kesamaan data antar *cluster* [9]. Dengan menggunakan teknik ini, kita dapat mengidentifikasi pola-pola yang tersembunyi dalam data, sehingga dapat membantu dalam pengambilan keputusan dan memahami karakteristik data yang sedang dianalisis [10]. Metode K-Means dapat diaplikasikan untuk melakukan pemetaan tingkat pengangguran di 35 Kabupaten/Kota di Jawa Tengah bertujuan guna mengidentifikasi lokasi pusat pengangguran dan pola sebaran *cluster* pengangguran [11]. Dalam kasus ini, pengelompokan tingkat pengangguran di Jawa Tengah akan dilakukan dengan menggunakan indikator jumlah penduduk, jumlah industri dan UMK [12]. Penelitian ini memanfaatkan indikator jumlah penduduk dan upah minimum kabupaten/kota selain jumlah pengangguran sebagai acuan utama.

Penelitian yang berhubungan dengan Algoritma K-Means telah dilakukan oleh Putri *et al.* (2022) mengenai algoritma pengelompokkan menggunakan *data mining* dengan Metode Spectral Clustering metode tersebut menganalisis mengenai studi kasus vaksinasi covid-19, penelitian tersebut menghasilkan 3 *cluster* yang evaluasi nya diukur menggunakan DBI dengan rata-rata sebesar 1.01422

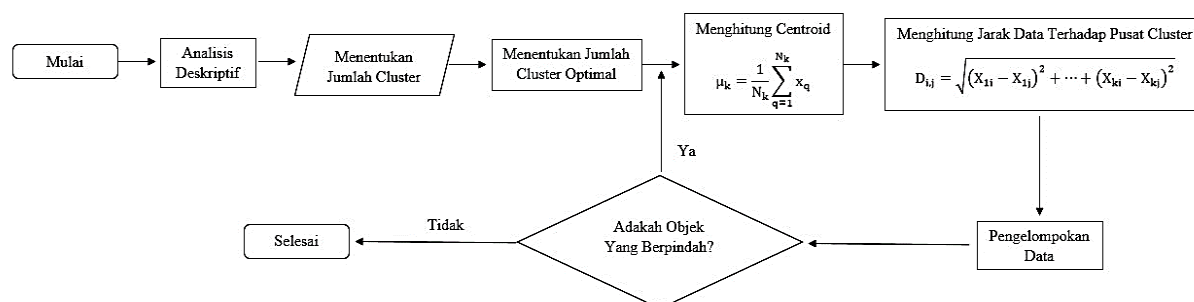
[13]. Penelitian selanjutnya dilakukan oleh Nainggolan *at al.* (2019) mengenai Peningkatan Performansi *K-Means* Cluster menggunakan *Sum of Square Error* (SSE) yang Dioptimalkan menggunakan Metode Elbow, penelitian tersebut mengenai tentang kelemahan algoritma *K-Means* menggunakan *Sum of Square Error* (SSE) untuk menghasilkan cluster yang lebih baik [14]. Penelitian selanjutnya dilakukan oleh Umargono *at al.* (2020) mengenai Optimasi *Clustering* menggunakan Metode Elbow dan Penentuan *Centroid* Awal, penelitian tersebut membahas tentang penentuan pusat cluster berdasarkan rata-rata membuat banyaknya iterasi dibutuhkan untuk mencapai keseragaman cluster 22.58% [15].

Berdasarkan konteks yang dijelaskan sebelumnya, penelitian ini guna memberikan gambaran yang akurat tentang kelompok daerah di Jawa Tengah yang mengalami tingkat pengangguran terbuka yang tinggi pada tahun 2021. Penelitian ini bertujuan menggali penyebab tingkat pengangguran terbuka dengan menganalisis lebih dalam hasil penelitian. Analisis *cluster* penelitian ini tidak hanya mengetahui jumlah cluster penyebab dari TPT, diharapkan dapat memberikan informasi yang bermanfaat untuk pemerintah dalam menggagaskan kebijakan yang sesuai untuk mengatasi isu pengangguran yang ada di Jawa Tengah. Dengan kata lain, diharapkan bahwa hasil artikel ini mampu memberikan kontribusi yang positif dalam mengatasi masalah pengangguran di Jawa Tengah.

METODE

Data Mining merupakan metode yang digunakan untuk menganalisis dan menggali informasi yang tersembunyi dari sekumpulan data yang telah diproses. Proses *data mining* ini bertujuan untuk mengungkap pola-pola yang ada dalam data tersebut. Dengan melakukan *data mining*, kita dapat menghasilkan informasi atau pengetahuan baru yang berharga dari sumber data yang sudah ada sebelumnya. Pengetahuan ini memiliki nilai yang sangat penting dan relevan untuk digunakan di masa depan [16]. Penerapan metode analisis dalam penelitian ini menggunakan analisis *cluster*. Analisis *cluster* merupakan salah satu teknik analisis multivariat yang menggunakan satu obyek kemudian menempatkan obyek tersebut ke dalam kelompok yang telah ditentukan berdasarkan karakteristik [17]. Dalam penelitian menggunakan metode alternatif yaitu metode non hierarki sebagai analisis data. Metode non hierarki merupakan salah satu metode pengelompokan obyek ke dalam beberapa kelompok berdasarkan jarak antar obyek [18].

Salah satu kelebihan metode pengelompokan adalah efisien dan kecepatannya, dan salah satu contohnya adalah Algoritma *K-Means*. *K-Means Clustering* sering digunakan dalam analisis data mining untuk menganalisis data operasional perusahaan yang tersimpan di dalam server basis data. Metode tersebut dapat diartikan guna mengelompokkan dan memetakan suatu data yang ada ke dalam beberapa kelompok *cluster* yang telah ditentukan berdasarkan karakteristik yang sama dengan kelompoknya dan berbeda dengan kelompok lainnya [19]. Tujuan dari metode ini guna mengelompokkan data menjadi *cluster*, data yang dikelompokkan tersebut memiliki karakteristik yang serupa satu sama lain. Pentingnya metode ini dilakukan pengelompokan tanpa adanya penentuan nilai acuan dalam prosesnya [20]. Langkah-langkah yang digunakan dalam penelitian ini meliputi [21]:



Gambar 1. Flowchart Algoritma K-Means

2.1 Jenis dan Sumber Data

Artikel ini menggunakan metode penelitian kuantitatif dalam pengolahan datanya. Penelitian kuantitatif merupakan jenis penelitian yang tidak terlalu memfokuskan pada pengukuran angka variabelnya. Desain penelitian kuantitatif ini dilakukan melalui beberapa tahapan diantaranya merumuskan masalah, mengkaji teori, merumuskan hipotesis, mengumpulkan data, menganalisis data, serta membuat kesimpulan dan saran.

Adapun sumber data penelitian ini berupa data sekunder yang diperoleh serta dikumpulkan dalam melakukan penelitian dari sumber yang sudah ada sebelumnya. Data yang digunakan berasal dari Badan Pusat Statistik (BPS) Jawa Tengah pada tahun 2021 yang dapat diakses di laman website (www.bps.go.id). Data setiap baris mewakili Kabupaten/Kota, sementara kolom – kolom $X_{1,1}$, $X_{2,1}$, $X_{3,1}$ menunjukkan nilai dari setiap variabel, disajikan pada tabel 1.

Tabel 1. Data Jumlah Penduduk, UMK, dan Jumlah Industri

Kabupaten/Kota	Variabel		
	X_1	X_2	X_3
Kab. Cilacap	$X_{1,1}$	$X_{2,1}$	$X_{3,1}$
Kab. Banyumas	$X_{1,2}$	$X_{2,2}$	$X_{3,2}$
Kab. Purbalingga	$X_{1,3}$	$X_{2,3}$	$X_{3,3}$
Kab. Banjarnegara	$X_{1,4}$	$X_{2,4}$	$X_{3,4}$
Kab. Kebumen	$X_{1,5}$	$X_{2,5}$	$X_{3,5}$
.	.	.	.
.	.	.	.
.	.	.	.
Kota Semarang	$X_{1,33}$	$X_{2,33}$	$X_{3,33}$
Kota Pekalongan	$X_{1,34}$	$X_{2,34}$	$X_{3,34}$
Kota Tegal	$X_{1,35}$	$X_{2,35}$	$X_{3,35}$

HASIL DAN PEMBAHASAN

3.1 Statistik Deskriptif

Karakteristik tingkat pengangguran terbuka di Provinsi Jawa Tengah, dapat dilihat pada tabel 2 dibawah ini :

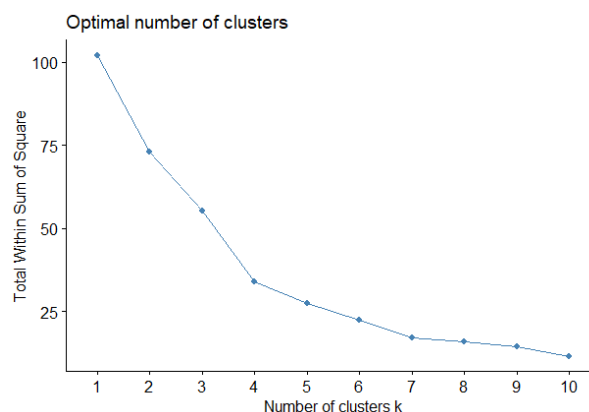
Tabel 2. Hasil Analisis Deskriptif dari 3 indikator

Jumlah Penduduk		UMK		Jumlah Industri	
Min.	: 121610	Min.	: 1805000	Min.	: 312
1st Qu.	: 829724	1st Qu.	: 1900200	1st Qu.	: 1697
Median	: 1025020	Median	: 1986450	Median	: 3911
Mean	: 1049786	Mean	: 2041504	Mean	: 5830
3rd Qu.	: 1318248	3rd Qu.	: 2104229	3rd Qu.	: 9088
Max.	: 1992685	Max.	: 2810025	Max.	: 18381

Berdasarkan tabel 2, diketahui bahwa pada Upah Minimum Kabupaten/Kota (UMK) memiliki nilai rata - rata tertinggi sebesar 2.104.229, yang menunjukkan rata - rata Upah Minimum Kabupaten/Kota di Provinsi Jawa Tengah. Rata - rata terendah pada Jumlah Industri sebesar 5.830 yang menunjukkan rata - rata banyak nya jumlah industri yang ada di Provinsi Jawa Tengah. Sedangkan pada jumlah penduduk memiliki rata - rata sebesar 1.049.786 juta jiwa penduduk di Provinsi Jawa Tengah.

3.2 Clustering K Means

Pengelompokkan K-Means dalam penentuan jumlah *cluster* optimal dengan menerapkan metode Elbow. Metode Elbow merupakan bagian dari metode *clustering* yang digunakan sebagai penentu jumlah *cluster* optimal dengan melihat titik jatuhnya yang berbentuk siku pada grafik. Hal ini dapat dilihat pada gambar 2.



Gambar 2. Grafik Metode Elbow

Grafik metode Elbow dilihat ketika titik *cluster* $k = 1$, nilai *total within_SS* yang paling tinggi. selanjutnya terus menurun drastis sampai pada jumlah *cluster* $k = 4$ terlihat membentuk seperti sebuah siku. Mulai jumlah *cluster* $k = 5$ sampai $k = 10$ terjadi mengalami penurunan secara stabil, namun nilai *total within_SS* sudah tidak signifikan. Sehingga dapat ditentukan jumlah *cluster* dari grafik metode Elbow sebanyak $k = 4$ [22].

Centroid atau titik pusat *cluster* yang diperoleh dari data yang diolah, menghasilkan *centroid* yang akan sangat berpengaruh terhadap hasil *cluster*. Jumlah kelompok atau *cluster* dapat dilihat dari jumlah *centroid*, jika ada empat *centroid* berarti ada empat kelompok yang dihasilkan. Berikut hasil *centroid* yang disajikan pada tabel 3.

Tabel 3. Pusat *Cluster* Pertama

cluster mean	Jumlah Penduduk	UMK	Jumlah Industri
1	0.2531124	-0.4498006	1.46329215
2	0.5341822	1.7771930	-0.09621854
3	0.4011814	-0.4297050	-0.55286417
4	-1.2871602	-0.1823115	-0.74474351

Nilai dari rata - rata pusat *cluster* yang terbentuk mengindikasikan bahwa **rata - rata** pada variabel **cluster** 1 merupakan titik pusat dari **cluster** 1, **rata - rata** dari **cluster** 2 merupakan titik pusat dari **cluster** 2, **rata - rata** dari **cluster** 3 merupakan titik pusat dari **cluster** 3, **rata - rata** dari **cluster** 4 merupakan titik pusat dari **cluster** 4 [23].

Perbedaan antara pusat **cluster** yang telah diolah dari data disajikan pada tabel 4.

Tabel 4. Jarak antar pusat **Cluster** Iterasi Pertama

	C1	C2	C3
C2	2.705441	0	0
C3	2.021686	1.717125	0
C4	2.733235	2.752732	2.257568

Berdasarkan tabel 4, selisih antara pusat cluster 1 ke cluster 2 sebesar 2.705441, cluster 1 ke cluster 3 sebesar 2.021686, Cluster 1 ke cluster 4 sebesar 2.733235, Cluster 2 ke cluster 3 sebesar 1.717125, Cluster 2 ke cluster 4 sebesar 2.752732, dan Cluster 3 ke cluster 4 sebesar 2.257568 [24]. Hasil pengelompokkan cluster pertama disajikan pada tabel 5.

Tabel 5. Pengelompokkan cluster pertama

Kab/Kota	C1	C2	C3	C4	Kab/Kota	C1	C2	C3	C4
Kab. Cilacap				*	Kab. Kudus				*
Kab. Banyumas	*				Kab. Jepara			*	
Kab. Purbalingga			*		Kab. Demak				*
Kab. Banjarnegara	*				Kab. Semarang				*
Kab. Kebumen			*		Kab. Temanggung		*		
Kab. Purworejo	*				Kab. Kendal				*
Kab. Wonosobo	*				Kab. Batang		*		
Kab. Magelang	*				Kab. Pekalongan	*			
Kab. Boyolali			*		Kab. Pemalang	*			
Kab. Klaten	*				Kab. Tegal			*	
Kab. Sukoharjo			*		Kab. Brebes			*	
Kab. Wonogiri			*		Kota Magelang		*		
Kab. Karanganyar			*		Kota Surakarta		*		
Kab. Sragen	*				Kota Salatiga		*		
Kab. Grobogan			*		Kota Semarang				*
Kab. Blora			*		Kota Pekalongan		*		
Kab. Rembang		*			Kota Tegal		*		
Kabupaten Pati			*						

Hasil cluster dari iterasi pertama selesai, dilanjutkan dengan pusat cluster baru untuk iterasi kedua menghitung nilai rata rata dari setiap cluster. Berikut nilai dari centroid untuk iterasi kedua pada tabel 6.

Tabel 6. Pusat Cluster Kedua

cluster mean	Jumlah Penduduk	UMK	Jumlah Industri
1	0.2531124	-0.4498006	1.46329215
2	0.5341822	1.7771930	-0.09621854
3	0.4011814	-0.4297050	-0.55286417
4	-1.2871602	-0.1823115	-0.74474351

Berdasarkan hasil centroid dari tabel 5, jika sampai pada tahap ini tidak terjadi perpindahan dan menghasilkan nilai yang sama maka perhitungan iterasi dihentikan [24]. Penelitian ini dilakukan sampai pada iterasi kedua, dapat dilihat pada tabel 7.

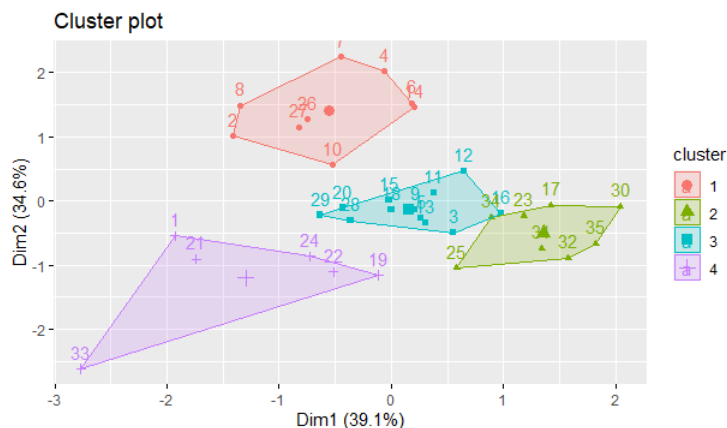
Tabel 7. Jarak antar pusat Cluster iterasi kedua

	C1	C2	C3
C2	2.705441	0	0
C3	2.021686	1.717125	0
C4	2.733235	2.752732	2.257568

Berdasarkan tabel 7, iterasi kedua menghasilkan nilai jarak antar pusat cluster yang sama dengan iterasi pertama serta tidak terjadi perpindahan. Hal ini menandakan cluster yang

terbentuk berhenti pada iterasi kedua artinya empat cluster data telah berkelompok menjadi cluster berdasarkan jarak rata-rata dari pusat cluster [25].

Setiap titik mewakili sebuah data yang di clusterkan dalam cluster plot warna untuk membedakan kelompok yang berbeda. Titik-titik pusat setiap kelompok ditandai dengan simbol lingkaran yang tersebar. Titik data yang berdekatan dan warna yang sama dalam plot cenderung menjadi bagian dari kelompok yang sama [26]. Plot K-Means memberikan gambaran visual tentang hasil clustering yang disajikan pada gambar 3.



Gambar 3. Plot K-Means

Berdasarkan gambar 3, hasil analisis cluster plot cluster 1 disimbolkan dengan warna merah, cluster 2 dengan warna hijau, cluster 3 dengan warna biru, dan cluster 4 dengan warna ungu. Dari gambar tersebut terlihat pada cluster 2 dan cluster 3 cenderung berdekatan dan berhimpit [26]. Banyaknya data yang tergabung dalam setiap kelompok dapat dilihat pada tabel 8.

Tabel 8. Jumlah Cluster Terbentuk

Cluster	Size	Kabupaten/Kota
Cluster 1	9	Kab. Banyumas, Kab. Banjarnegara, Kab. Purworejo, Kab. Wonosobo, Kab. Magelang, Kab. Klaten, Kab. Sragen, Kab. Pekalongan, Kab. Pemalang
Cluster 2	6	Kab. Cilacap, Kab. Kudus, Kab. Demak, Kab. Semarang, Kab. Kendal, Kota Semarang
Cluster 3	12	Kab. Purbalingga, Kab. Kebumen, Kab. Boyolali, Kab. Sukoharjo, Kab. Wonogiri, Kab. Karanganyar, Kab. Grobogan, Kab. Blora, Kab. Pati, Kab. Jepara, Kab. Tegal, Kab. Brebes,
Cluster 4	8	Kab. Rembang, Kab. Temanggung, Kab. Batang, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Pekalongan, Kota Tegal

Jumlah cluster yang terbentuk sebanyak $k = 4$. Jumlah anggota dari masing - masing cluster adalah 9, 8, 12, 6 yang menunjukkan bahwa nilai cluster dari K-Means terbentuk dari cluster 1 sebanyak 9 Kabupaten/Kota, cluster 2 sebanyak 8 Kabupaten/Kota, cluster 3 sebanyak 12 Kabupaten/Kota, dan Cluster 4 sebanyak 6 Kabupaten/Kota. Selanjutnya diperoleh nilai within Cluster Sum Of Square yang disajikan pada tabel 8.

Tabel 8. Akurasi Clustering K-Means

Within Cluster Sum Of Square			
C1	C2	C3	C3
8.703134	5.506497	8.991976	10.639443

Berdasarkan tabel 8, Hasil akurasi cluster yang terbentuk didapatkan dari nilai *between_SS* (*between Sum of Square*) dibagi nilai *total_SS* (*total Sum Of Square*) sehingga mendapatkan nilai sebesar 66.8%, yang mempresentasikan seberapa berbeda nilai antar cluster dengan cluster lainnya. Akurasi cluster memberikan gambaran baiknya kualitas pengelompokan, secara keseluruhan semakin tinggi nilai SSE maka hasil clustering semakin baik kualitas clusternya [27].

KESIMPULAN

Berdasarkan artikel yang dianalisis menggunakan metode k-means, dapat disimpulkan bahwa :

1. Hasil dari analisis deskriptif yang dilakukan menghasilkan nilai minimum dari jumlah penduduk di Provinsi Jawa Tengah sebesar 122.1 dan nilai maksimum sebesar 1.992.685. Rata – rata jumlah penduduk sebesar 1.049.786. Pada indikator UMK memiliki nilai minimum sebesar 1.805.000 dan nilai maksimum sebesar 2.810.025. Rata – rata UMK sebesar 2.041.504. Pada indikator jumlah industri memiliki nilai minimum sebesar 312 dan nilai maksimum sebesar 18.381. Rata – rata jumlah industri sebesar 5.830.
2. Hasil penelitian analisis cluster K-Means dengan 4 cluster melalui proses sebanyak 2 iterasi. Penentuan centroid menghasilkan nilai jarak pusat cluster iterasi pertama, hasil cluster sebanyak 9 Kabupaten/Kota, 8 Kabupaten/Kota, 12 Kabupaten/Kota, dan 6 Kabupaten/Kota. Dilanjutkan penentuan centroid baru menghasilkan nilai jarak pusat cluster iterasi kedua, dengan hasil cluster sama seperti iterasi pertama dan tidak terjadi perubahan sehingga berhenti di iterasi kedua.
3. Hasil akurasi yang terbentuk didapatkan nilai sebesar 66.8%. Hal ini menunjukkan bahwa nilai akurasi cukup baik, sehingga diharapkan pada penelitian selanjutnya untuk melakukan optimalisasi agar menghasilkan nilai akurasi yang lebih baik lagi.

DAFTAR PUSTAKA

- [1.] Hegelund, E., & Taalbi, J. (2023). What determines unemployment in the long run? Band spectrum regression on ten countries 1913–2016. *Structural Change and Economic Dynamics*, 64, 144–167. <https://doi.org/10.1016/j.strueco.2022.11.009>
- [2.] Utami, T. W., Rohman, A., & Prahutama, A. (2017). Pemodelan Regresi Berganda Dan Geographically Weighted Regression Pada Tingkat Pengangguran Terbuka Di Jawa Tengah. *Media Statistika*, 9(2), 133–147. <https://doi.org/10.14710/medstat.9.2.133-147>
- [3.] Akramunnisa, A., & Fajriani, F. (2020). K-Means Clustering Analysis pada Persebaran Tingkat Pengangguran Kabupaten/Kota di Sulawesi Selatan. *Jurnal Varian*, 3(2), 103–112. <https://doi.org/10.30812/varian.v3i2.652>
- [4.] Hendri Doni, A., Alfiona, F., Andespa, W., Islam, E., & Islam Negeri Sjech Djamil Djambek Bukittinggi, U. M. (2022–2023). Pengangguran Dalam Perspektif Ekonomi Islam Dan Konvensional. *Manajemen dan Syariah JIEMAS*, 2(1). <https://doi.org/10.55883/jiemas.v2i1>

- [5.] Damayanti, A., & Fisabilillah Perdini, W. L. (2022). Pengaruh Tingkat Pengangguran Terbuka dan Pertumbuhan Penduduk Terhadap Tingkat Kemiskinan Di Kota Malang. *Independent: Journal of Economics*, 2(1), 1–15. Link
- [6.] Pasuria, S., & Triwahyuningtyas, N. (2022). Pengaruh Angkatan Kerja, Pendidikan, Upah Minimum, Dan Produk Domestik Bruto Terhadap Pengangguran Di Indonesia. *Sibatik Journal: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, dan Pendidikan*, 1(6), 795–808. <https://doi.org/10.54443/sibatik.v1i6.94>
- [7.] Berliani, K. (2021). Pengaruh Tingkat Pengangguran, Tingkat Pendidikan dan Laju Pertumbuhan Penduduk Terhadap Tingkat Kemiskinan Penduduk Provinsi Jawa Barat Tahun 2015-2020. *Syntax Literate; Jurnal Ilmiah Indonesia*, 6(2), 872. <https://doi.org/10.36418/syntax-literate.v6i2.2244>
- [8.] Aldino, A. A., Darwis, D., Prastowo, A. T., & Sujana, C. (2021). Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency. *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/1751/1/012038>
- [9.] Aytaç, E. (2020). Unsupervised learning approach in defining the similarity of catchments: Hydrological response unit based k-means clustering, a demonstration on Western Black Sea Region of Turkey. *International Soil and Water Conservation Research*, 8(3), 321–331. <https://doi.org/10.1016/j.iswcr.2020.05.002>
- [10.] Kurnia, F., Fahmi, I., Wahyudi, E., Mige, G. E., & Sultan Syarif Kasim Fakultas Sains, U. R. (2019). Penerapan Algoritma K-Means Untuk Pengelompokan Diagnosa Penyakit Mata Berdasarkan Rentang Usia. *Jurnal SPEKTRO*, 2(1).
- [11.] Talasbek, A., Serek, A., Zhaparov, M., Moo-Yoo, S., Kim, Y. K., & Jeong, G. H. (2020). Personality classification experiment by applying k-means clustering. *International Journal of Emerging Technologies in Learning*, 15(16), 162–177. <https://doi.org/10.3991/ijet.v15i16.15049>
- [12.] Hartono, R., Busari, A., & Awaluddin, M. (2018). Pengaruh produk domestik regional bruto (pdrb) dan upah minimum kota (umk) terhadap penyerapan tenaga kerja. *Inovasi*, 14(1), 36–43. Link
- [13.] Winadya Putri, M., Manfaati Nur, I., & Wasono, R. (2022). Implementasi Spectral Clustering Algorithm Untuk Pengelompokan Sasaran Vaksinasi Covid-19 Di Indonesia. *Jurnal Statistika*, 10(1).
- [14.] Nainggolan, R., Perangin-Angin, R., Simarmata, E., & Tarigan, A. F. (2019). Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method. *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/1361/1/012015>
- [15.] Umargono, E., Suseno, J. E., & Gunawan, V. (2020). K-Means Clustering Optimization Using the Elbow Method and Early Centroid Determination Based on Mean and Median Formula. *Advances in Social Science, Education and Humanities Research*, 474.
- [16.] Terpadu, J. T., Nabarian, T., Ganiardi, M. A., & Malik, R. F. (2020). Implementasi Metode Hybrid Fuzzy C-Means Dan Fuzzy Swarm Untuk Pengelompokan Data Benang Perusahaan Tekstil. *Journal of Integrated Technology*, 6(1), 39–45. Link
- [17.] Nofiantika Rosulya Putri, C., Kurniawati, E., Manfaati Nur, I., & Wasono, R. (2021). Complete Linkage Hierarchical Clustering Clustering and Mapping Sub-Districts in Kendal
- [18.] Zaifullah, Z., & Yulianto, T. (2022). Analisis Cluster Untuk Pengelompokan Prestasi Mahasiswa Angkatan 2013 Fakultas MIPA Universitas Islam Madura. *Zeta - Math Journal*, 7(1), 1–10. doi: 10.31102/zeta.2022.7.1.1-10.
- [19.] Mashur, M. I., & Salim, Y. (2022). Analisis Performa Metode Cluster K-Means pada Dataset Ocular Disease Recognition. *Indonesian Journal of Data and Science (IJODAS)*, 3(1), 35–46.

- [20.] Irawan, Y., Hang, S., & Pekanbaru, T. (2019). Implementation Of Data Mining For Determining Majors Using K-Means Algorithm In Students Of Sma Negeri 1 Pangkalan Kerinci. *Journal of Applied Engineering and Technological Science*, 1(1), 17-29.
- [21.] Hidayatur Rifa, I., & Pratiwi, H. (2020). Clustering Of Earthquake Risk In Indonesia Using K-Medoids And K-Means Algorithms. *Media Statistika*, 13(2), 194–205. doi: 10.14710/medstat.13.1.194-205.
- [22.] Yuliana Sari, R., Oktavianto, H., & Wahyu Sulisty, H. (2022). Algoritma K-Means Dengan Metode Elbow Untuk Mengelompokkan Kabupaten/Kota Di Jawa Tengah Berdasarkan Komponen Pembentuk Indeks Pembangunan Manusia. *Jurnal Smart Teknologi*, 3(2), 104–108. <http://jurnal.unmuhember.ac.id/index.php/JST>
- [23.] Triyani, E., Agustian Hudjimartsu, S., & Primasari, D. (2022). Spasial Clustering Potensi Peternakan Unggas Dengan Metode K-Means Berbasis Webgis. *Infotech Jurnal*, 8(2). doi: 10.31949/infotech.v8i2.2627.
- [24.] Havaluddin, H., Patandianan, S. J., Putra, G. M., Puspitasari, N., & Pakpahan, H. S. (2021). Implementasi Metode K-Means Untuk Pengelompokkan Rekomendasi Tugas Akhir. *Informatika Mulawarman: Jurnal Ilmiah Ilmu Komputer*, 16(1), 13. doi: 10.30872/jim.v16i1.5182.
- [25.] Fahmi, M., Adi, H., & Susetyo, Y. A. (2020). Pemetaan Potensi Bencana di Jawa Tengah Menggunakan Google Maps API dan KML dengan Metode K-Means. *Indonesian Journal of Computing and Modeling*, 3(1), 43-51.
- [26.] Nurhidayati, M., & Khasanah, N. (2021). Penggunaan Metode K-Means Cluster Untuk Mengklasifikasikan Kemampuan 4C Mahasiswa. *Jurnal Ilmiah Matematika Dan Terapan*, 18(2), 160–169. doi: 10.22487/2540766x.2021.v18.i2.15615.
- [27.] Ayu, D., Dewi, I. C., & Pramita, K. (2019). Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *Jurnal Matrix*, 9(3).