

Penerapan Metode K-Means pada Pengelompokan Persebaran Kasus Diabetes Mellitus di Kecamatan Kersana

Muhimatul Ifadah¹, Agyztia Premana, M. Kom², Nur Ariesanto Ramdhan, M. Kom³

^{1, 2, 3}Program Studi Teknik Informatika, Universitas Muhadi Setiabudi, Indonesia

Email author: muhimatulifadah15@gmail.com¹, a.premana@umus.ac.id²,
ariesantoramdhan@gmail.com³

ABSTRAK

Diabetes Mellitus merupakan salah satu penyakit tidak menular yang memerlukan perhatian dalam upaya pencegahan dan pengendalian kesehatan masyarakat. Persebaran kasus pada setiap desa memiliki tingkat yang berbeda sehingga diperlukan metode analisis untuk mengidentifikasi kelompok wilayah berdasarkan tingkat kasus. Penelitian ini bertujuan menerapkan metode K-Means Clustering untuk mengelompokkan persebaran kasus Diabetes Mellitus pada 13 desa di Kecamatan Kersana. Data penelitian terdiri atas 365 kasus, yaitu 82 kasus laki-laki dan 283 kasus perempuan. Penelitian menggunakan pendekatan kuantitatif deskriptif dengan jumlah kasus Diabetes Mellitus sebagai variabel utama clustering, sedangkan data jenis kelamin digunakan sebagai informasi pendukung. Pengelompokan dilakukan menjadi tiga cluster, yaitu rendah, sedang, dan tinggi. Hasil pengelompokan menunjukkan bahwa cluster rendah terdiri atas 9 desa dengan rata-rata 15,33 kasus, cluster sedang terdiri atas 2 desa dengan rata-rata 36 kasus, dan cluster tinggi terdiri atas 2 desa dengan rata-rata 77,50 kasus. Desa Limbangan dan Kradenan termasuk cluster tinggi, Cikandang dan Kersana termasuk cluster sedang, sedangkan Cigedog, Kubangpari, Ciampel, Jagapura, Sutamaja, Kemukten, Sindangjaya, Pande, dan Keramatsampang termasuk cluster rendah. Evaluasi menggunakan Silhouette Score menghasilkan nilai 0,690 yang menunjukkan struktur cluster cukup baik. Hasil penelitian menunjukkan bahwa K-Means dapat membantu mengidentifikasi kelompok wilayah berdasarkan tingkat kasus Diabetes Mellitus sebagai informasi pendukung dalam menentukan prioritas pemantauan dan perencanaan program kesehatan.

Kata Kunci: *Diabetes Mellitus, K-Means, Clustering, Pengelompokan Wilayah, Kecamatan Kersana*

ABSTRACT

Diabetes Mellitus is a non-communicable disease that requires attention in public health prevention and control efforts. The distribution of cases varies across villages, requiring an analytical method to identify regional groups based on case levels. This study aims to apply the K-Means Clustering method to classify the distribution of Diabetes Mellitus cases in 13 villages in Kersana District. The data consist of 365 cases, comprising 82 male cases and 283 female cases. This study uses a descriptive quantitative approach, with the total number of Diabetes Mellitus cases as the main clustering variable, while gender data are used as supporting descriptive information. The clustering process divides the villages into three clusters: low, medium, and high. The results show that the low-level cluster consists of 9 villages with an average of 15.33 cases, the medium-level cluster consists of 2 villages with an average of 36 cases, and the high-level cluster consists of 2 villages with an average of 77.50 cases. Limbangan and Kradenan are categorized as high-level areas, Cikandang and Kersana as medium-level areas, while Cigedog, Kubangpari, Ciampel, Jagapura, Sutamaja, Kemukten, Sindangjaya, Pande, and Keramatsampang are categorized as low-level areas. The evaluation using the Silhouette Score resulted in a value of 0.690, indicating a reasonably good cluster structure. The results demonstrate that K-Means can help identify regional groups based on Diabetes Mellitus case levels and provide supporting information for prioritizing health monitoring and planning health programs.

Keywords: *Diabetes Mellitus, K-Means, Clustering, Regional Grouping, Kersana District*

1. PENDAHULUAN

Diabetes Mellitus merupakan salah satu penyakit tidak menular yang menjadi perhatian dalam bidang kesehatan karena berkaitan dengan gangguan metabolisme dan dapat menimbulkan berbagai komplikasi apabila tidak dikelola dengan baik. (Aminuddin et al., 2023) Peningkatan jumlah penderita Diabetes Mellitus dapat memberikan dampak terhadap kualitas hidup masyarakat serta meningkatkan kebutuhan terhadap pelayanan kesehatan. Oleh karena itu, pemantauan persebaran kasus pada tingkat wilayah menjadi salah satu bagian penting dalam mendukung perencanaan program kesehatan.

Informasi mengenai jumlah kasus Diabetes Mellitus pada suatu wilayah tidak hanya dapat digunakan untuk mengetahui besarnya permasalahan kesehatan, tetapi juga dapat dimanfaatkan untuk mengidentifikasi wilayah yang memiliki tingkat kasus relatif tinggi, sedang, maupun rendah. Pengelompokan wilayah tersebut dapat membantu pihak terkait dalam menentukan prioritas pemantauan dan menyusun strategi intervensi kesehatan yang lebih terarah.

Kecamatan Kersana merupakan salah satu wilayah administratif di Kabupaten Brebes yang terdiri atas beberapa desa. Berdasarkan data yang digunakan dalam penelitian ini, terdapat 13 desa dengan total 365 kasus Diabetes Mellitus. Jumlah kasus pada setiap desa menunjukkan variasi yang cukup besar. Desa Limbangan memiliki 79 kasus dan Desa Kradenan memiliki 76 kasus, sedangkan Desa Sindangjaya memiliki 8 kasus. Perbedaan tersebut menunjukkan adanya variasi tingkat kasus antarwilayah yang perlu dianalisis secara sistematis.

Data yang tersedia juga menunjukkan perbedaan berdasarkan jenis kelamin. Dari total 365 kasus, sebanyak 82 kasus atau sekitar 22,47% merupakan kasus laki-laki, sedangkan 283 kasus atau sekitar 77,53% merupakan kasus

perempuan. Informasi tersebut dapat digunakan untuk memberikan gambaran karakteristik data, meskipun penelitian ini memfokuskan proses pengelompokan pada jumlah kasus Diabetes Mellitus di masing-masing desa.

Salah satu metode yang dapat digunakan untuk melakukan pengelompokan data adalah *K-Means Clustering*. *K-Means* merupakan metode *unsupervised learning* yang mengelompokkan data berdasarkan kemiripan karakteristik ke dalam sejumlah kelompok atau cluster. Algoritma ini bekerja dengan menentukan sejumlah centroid awal, menghitung jarak setiap data terhadap centroid, mengalokasikan data ke centroid terdekat, (Irawan et al., 2025) kemudian memperbarui centroid hingga kondisi konvergen tercapai.

Penerapan *K-Means* pada data kasus Diabetes Mellitus dapat menghasilkan kelompok wilayah berdasarkan kemiripan tingkat jumlah kasus. (Mafela et al., 2025) Hasil pengelompokan tersebut diharapkan dapat memberikan informasi yang lebih mudah dipahami mengenai wilayah dengan tingkat kasus rendah, sedang, dan tinggi.

Penelitian ini memiliki perbedaan dengan analisis statistik deskriptif biasa karena tidak hanya menyajikan jumlah kasus pada setiap desa, tetapi juga melakukan proses pengelompokan secara komputasional menggunakan algoritma *K-Means*. Dengan demikian, hasil penelitian dapat memberikan pendekatan berbasis data dalam mengidentifikasi kelompok wilayah berdasarkan tingkat kasus Diabetes Mellitus.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menerapkan metode *K-Means* dalam mengelompokkan persebaran kasus Diabetes Mellitus pada desa-desa di Kecamatan Kersana. Hasil pengelompokan diharapkan dapat menjadi informasi pendukung bagi pihak terkait dalam memahami pola tingkat kasus dan menentukan wilayah yang perlu

mendapatkan perhatian lebih dalam kegiatan pemantauan kesehatan.

dilakukan berdasarkan karakteristik data secara otomatis.

2. LANDASAN TEORI

2.1 Diabetes Mellitus

Diabetes Mellitus merupakan penyakit metabolik kronis yang ditandai dengan peningkatan kadar glukosa dalam darah akibat gangguan produksi insulin, kerja insulin, atau kombinasi keduanya. Kondisi tersebut dapat berlangsung dalam jangka panjang dan berpotensi menyebabkan komplikasi pada berbagai organ tubuh. (Magelang, 2025)

Pengendalian Diabetes Mellitus memerlukan pendekatan yang berkelanjutan, baik melalui perubahan gaya hidup, pemeriksaan kesehatan, maupun pengelolaan medis. (Korespondensi, 2025) Oleh karena itu, informasi mengenai persebaran kasus pada tingkat wilayah dapat membantu dalam proses perencanaan program kesehatan.

Dalam konteks penelitian ini, data Diabetes Mellitus digunakan sebagai objek analisis untuk mengelompokkan desa berdasarkan jumlah kasus yang tercatat. Pengelompokan tersebut tidak dimaksudkan untuk menentukan diagnosis individu, tetapi untuk melihat tingkat kasus secara agregat pada wilayah desa.

2.2 Data Mining

Data mining merupakan proses menemukan pola, informasi, atau pengetahuan yang bermanfaat dari sekumpulan data. (Luhur et al., 2020) Proses tersebut dapat dilakukan dengan berbagai metode, seperti klasifikasi, regresi, asosiasi, dan clustering.

Dalam penelitian ini, metode *data mining* yang digunakan adalah clustering karena data yang tersedia belum memiliki label kelompok rendah, sedang, atau tinggi. Dengan menggunakan metode clustering, pengelompokan dapat

2.3 K-Means Clustering

K-Means merupakan salah satu algoritma clustering yang bertujuan membagi data ke dalam sejumlah cluster berdasarkan kedekatan terhadap pusat cluster atau centroid. (Norshahlan et al., 2023)

Secara umum, tahapan K-Means adalah sebagai berikut:

1. Menentukan jumlah cluster yang akan digunakan.
2. Menentukan centroid awal.
3. Menghitung jarak setiap data terhadap masing-masing centroid.
4. Menentukan cluster berdasarkan jarak terdekat.
5. Menghitung centroid baru berdasarkan rata-rata data dalam setiap cluster.
6. Mengulangi proses perhitungan jarak dan pembaruan centroid.
7. Menghentikan proses apabila centroid tidak mengalami perubahan atau kondisi konvergensi telah tercapai.

Perhitungan jarak Euclidean dapat dituliskan sebagai:

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

Keterangan:

- $d(x_i, c_j)$ = jarak data ke-i terhadap centroid ke-j;
- x_{ik} = nilai data ke-i pada atribut ke-k;
- c_{jk} = nilai centroid ke-j pada atribut ke-k.

Centroid baru dihitung menggunakan:

$$c_j = \frac{1}{N_j} \sum_{i=1}^{N_j} x_i$$

Keterangan:

- C_j = centroid baru cluster ke- j ;
- N_j = jumlah data dalam cluster ke- j ;
- x_i = data anggota cluster.

2.4 Clustering Wilayah Kasus Diabetes Mellitus

Clustering wilayah merupakan proses pengelompokan wilayah berdasarkan karakteristik tertentu. (Suryani et al., 2026) Dalam penelitian ini, karakteristik yang digunakan adalah jumlah kasus Diabetes Mellitus pada setiap desa.

Penggunaan clustering memungkinkan wilayah yang memiliki jumlah kasus relatif sama berada dalam satu kelompok. Hasil pengelompokan kemudian dapat diberi interpretasi berupa kategori rendah, sedang, dan tinggi berdasarkan nilai centroid.

2.5 Evaluasi Clustering

Kualitas hasil clustering dapat dievaluasi menggunakan *Silhouette Score*. (Ishak, 2024) Nilai *Silhouette Score* berada pada rentang -1 hingga 1. Nilai yang semakin mendekati 1 menunjukkan bahwa data dalam cluster cenderung memiliki kemiripan yang tinggi dan memiliki pemisahan yang lebih baik dari cluster lainnya.

Dalam penelitian ini, *Silhouette Score* digunakan sebagai evaluasi tambahan untuk mengetahui kualitas struktur cluster yang dihasilkan oleh algoritma K-Means.

3. METODE PENELITIAN

3.1 Jenis Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif. Pendekatan kuantitatif digunakan karena data penelitian berupa angka jumlah kasus Diabetes Mellitus pada masing-masing desa. Sementara itu, pendekatan deskriptif digunakan untuk menggambarkan hasil pengelompokan wilayah berdasarkan tingkat kasus.

3.2 Lokasi Penelitian

Penelitian dilakukan pada wilayah Kecamatan Kersana, Kabupaten Brebes, Jawa Tengah. Objek penelitian adalah data kasus Diabetes Mellitus yang tercatat pada 13 desa.

3.3 Data Penelitian

Data yang digunakan terdiri atas:

1. Nama desa.
2. Jumlah kasus Diabetes Mellitus.
3. Jumlah kasus laki-laki.
4. Jumlah kasus perempuan.

Data terdiri atas 13 baris observasi, dengan total 365 kasus.

3.4 Preprocessing Data

Tahap preprocessing dilakukan sebelum proses clustering. Tahapan tersebut meliputi:

1. Memeriksa kelengkapan data.
2. Memastikan tidak terdapat data kosong pada atribut utama.
3. Memeriksa kesesuaian nama desa.
4. Memastikan nilai jumlah kasus berupa data numerik.
5. Memastikan konsistensi jumlah kasus dengan data jenis kelamin.
6. Menggunakan jumlah kasus sebagai variabel utama clustering.

Berdasarkan data penelitian:

$$\text{Laki} - \text{laki} + \text{Perempuan} = \text{Jumlah Kasus}$$

Sebagai contoh:

$$6 + 14 = 20$$

pada Desa Cigedog.

Validasi tersebut dilakukan pada seluruh data sehingga data jumlah kasus konsisten dengan penjumlahan data laki-laki dan perempuan.

3.5 Variabel Penelitian

Variabel utama yang digunakan dalam proses K-Means adalah:

$$X = \text{Jumlah Kasus}$$

Pemilihan satu variabel utama dilakukan karena jumlah kasus merupakan indikator langsung yang digunakan untuk menentukan tingkat persebaran kasus Diabetes Mellitus antarwilayah.

Data laki-laki dan perempuan tidak digunakan sebagai variabel clustering utama karena keduanya membentuk jumlah kasus:

$$\text{Jumlah Kasus} = \text{Laki} - \text{laki} + \text{Perempuan}$$

Apabila ketiga variabel digunakan secara bersamaan, informasi yang sama akan masuk ke model lebih dari satu kali sehingga dapat memberikan bobot yang tidak proporsional terhadap jumlah kasus.

Data jenis kelamin tetap digunakan sebagai analisis deskriptif untuk menggambarkan karakteristik kasus Diabetes Mellitus.

3.6 Penentuan Jumlah Cluster

Penelitian ini menggunakan tiga cluster, yaitu:

- Cluster 1: Rendah
- Cluster 2: Sedang
- Cluster 3: Tinggi

Penentuan tiga cluster disesuaikan dengan tujuan penelitian untuk menghasilkan klasifikasi wilayah yang mudah diinterpretasikan.

4. HASIL DAN PEMBAHASAN

4.1 Analisis Deskriptif Data

Data penelitian terdiri atas 13 desa dengan total 365 kasus Diabetes Mellitus.

Tabel 4.1 data kasus

desa	Jumlah kasus	Laki-laki	Perempuan
Cigedog	20	6	14
Cikandang	31	13	18
Kersana	41	6	35
Kubangpari	13	5	8
Limbangan	79	12	67
Kradenan	76	8	68
Ciampel	22	9	13

Jagapura	22	3	19
Sutamaja	17	8	9
Kemukten	12	2	10
Sindangjaya	8	1	7
Pande	11	7	4
Keramatsampang	13	2	11
	365	82	283

Berdasarkan data tersebut, jumlah kasus tertinggi terdapat di Desa Limbangan sebanyak 79 kasus, diikuti oleh Desa Kradenan sebanyak 76 kasus. Sementara itu, jumlah kasus terendah terdapat di Desa Sindangjaya sebanyak 8 kasus.

Dari aspek jenis kelamin, kasus perempuan berjumlah 283 kasus atau sekitar 77,53%, sedangkan laki-laki sebanyak 82 kasus atau sekitar 22,47%. Hasil ini menunjukkan bahwa data kasus pada penelitian didominasi oleh kasus perempuan. Namun, perbedaan jumlah kasus berdasarkan jenis kelamin tidak menjadi dasar utama pengelompokan K-Means karena fokus penelitian adalah pengelompokan wilayah berdasarkan jumlah kasus.

4.2 Implementasi K-Means

4.2.1 Penentuan Jumlah Cluster

Jumlah cluster yang digunakan adalah:

$$K = 3$$

Ketiga cluster tersebut diinterpretasikan sebagai:

- Cluster rendah
- Cluster sedang
- Cluster tinggi

4.2.2 Centroid Awal

Dalam implementasi menggunakan algoritma K-Means, centroid awal dapat ditentukan secara otomatis menggunakan mekanisme inisialisasi K-Means++. (Studi et al., 2026) Dengan proses tersebut, algoritma mencari pusat cluster awal yang dapat membantu memperoleh pengelompokan yang stabil.

Karena hanya terdapat satu atribut numerik utama, yaitu jumlah kasus, proses clustering dapat dipahami sebagai pengelompokan data berdasarkan kedekatan nilai jumlah kasus terhadap pusat cluster.

4.2.3 Centroid Akhir

Setelah proses iterasi selesai, diperoleh centroid akhir sebagai berikut:

Cluster Centroid Interpretasi

Cluster 1 15,33 Rendah

Cluster 2 36,00 Sedang

Cluster 3 77,50 Tinggi

Centroid tersebut menunjukkan karakteristik rata-rata jumlah kasus pada setiap kelompok.

Cluster rendah memiliki nilai pusat 15,33 kasus. Cluster sedang memiliki nilai pusat 36 kasus. Sementara itu, cluster tinggi memiliki nilai pusat 77,50 kasus.

4.3 Hasil Pengelompokan Wilayah

Hasil akhir proses K-Means ditunjukkan pada tabel berikut.

Tabel 4.2 Hasil Cluster

No	Desa	Jumlah Kasus	Cluster	Kategori
1	Cigedog	20	C1	Rendah
2	Cikandang	31	C2	Sedang
3	Kersana	41	C2	Sedang
4	Kubangpari	13	C1	Rendah
5	Limbangan	79	C3	Tinggi
6	Kradenan	76	C3	Tinggi
7	Ciampel	22	C1	Rendah
8	Jagapura	22	C1	Rendah
9	Sutamaja	17	C1	Rendah

10	Kemukten	12	C1	Rendah
11	Sindangjaya	8	C1	Rendah
12	Pande	11	C1	Rendah
13	Keramatsampang	13	C1	Rendah

Berdasarkan hasil tersebut, diperoleh tiga kelompok wilayah.

1. Cluster Rendah

Cluster rendah terdiri atas:

- Cigedog;
- Kubangpari;
- Ciampel;
- Jagapura;
- Sutamaja;
- Kemukten;
- Sindangjaya;
- Pande;
- Keramatsampang.

Jumlah anggota cluster adalah 9 desa dengan rentang jumlah kasus antara 8 hingga 22 kasus. Nilai centroid cluster sebesar 15,33 kasus.

Cluster ini merupakan kelompok dengan jumlah kasus relatif rendah dibandingkan dengan kelompok lainnya.

2. Cluster Sedang

Cluster sedang terdiri atas:

- Cikandang;
- Kersana.

Jumlah anggota cluster sebanyak 2 desa dengan jumlah kasus masing-masing 31 dan 41 kasus. Nilai centroid cluster adalah:

$$\frac{31 + 41}{2} = 36$$

Dengan demikian, centroid cluster sedang adalah 36 kasus.

3. Cluster Tinggi

Cluster tinggi terdiri atas:

- Limbangan;
- Kradenan.

Jumlah kasus masing-masing desa adalah 79 dan 76 kasus. Centroid cluster dihitung sebagai:

$$\frac{79 + 76}{2} = 77,5$$

Dengan demikian, centroid cluster tinggi adalah 77,50 kasus.

Kedua desa tersebut memiliki jumlah kasus yang jauh lebih tinggi dibandingkan dengan sebagian besar desa lainnya sehingga menjadi wilayah yang perlu mendapatkan perhatian lebih dalam pemantauan kasus.

4.4 Distribusi Jumlah Desa pada Setiap Cluster

Tabel 4.3 hasil pengelompokan

Kategori	Jumlah Desa	Persentase
Rendah	9	69,23%
Sedang	2	15,38%
Tinggi	2	15,38%
Total	13	100%

Hasil pengelompokan menunjukkan bahwa sebagian besar wilayah, yaitu 9 dari 13 desa atau 69,23%, termasuk dalam cluster rendah. Sebanyak 2 desa atau 15,38% masuk dalam cluster sedang, dan 2 desa atau 15,38% termasuk cluster tinggi.

Meskipun jumlah desa dalam cluster tinggi hanya dua desa, kedua wilayah tersebut memiliki jumlah kasus yang sangat besar. Limbangan menyumbang 79 kasus dan Kradenan 76 kasus. Secara bersama-sama, kedua desa tersebut memiliki:

$$79 + 76 = 155$$

kasus atau sekitar:

$$\frac{155}{365} \times 100\% = 42,47\%$$

dari keseluruhan kasus.

Temuan ini penting karena menunjukkan bahwa dua desa dalam cluster tinggi menyumbang sekitar 42,47% dari seluruh kasus, meskipun hanya mencakup 15,38% dari jumlah desa yang dianalisis.

4.5 Evaluasi Hasil Clustering

Evaluasi menggunakan *Silhouette Score* menghasilkan nilai:

$$\text{Silhouette Score} = 0,690$$

Nilai tersebut menunjukkan bahwa hasil pengelompokan memiliki struktur cluster yang cukup baik. Data pada masing-masing kelompok cenderung memiliki kedekatan karakteristik berdasarkan jumlah kasus, sementara pemisahan antarcluster relatif jelas.

Cluster rendah memiliki rentang 8–22 kasus, cluster sedang berada pada 31–41 kasus, sedangkan cluster tinggi berada pada 76–79 kasus. Terdapat jarak yang cukup besar antara cluster sedang dan tinggi, sehingga kedua kelompok tersebut dapat dibedakan dengan relatif jelas.

Hasil evaluasi ini memperkuat hasil pengelompokan yang diperoleh dari proses K-Means.

4.6 Pembahasan

Hasil penelitian menunjukkan bahwa metode K-Means dapat diterapkan untuk mengelompokkan wilayah berdasarkan jumlah kasus Diabetes Mellitus. Dengan menggunakan tiga cluster, wilayah Kecamatan Kersana dapat dibagi menjadi kelompok rendah, sedang, dan tinggi.

Cluster rendah merupakan kelompok terbesar dengan 9 desa. Hal ini menunjukkan bahwa sebagian besar desa dalam dataset memiliki jumlah kasus yang relatif rendah, yaitu antara 8 hingga 22 kasus.

Cluster sedang terdiri atas Cikandang dan Kersana. Kedua desa tersebut memiliki jumlah kasus masing-masing 31 dan 41 kasus. Nilai tersebut berada di atas kelompok rendah tetapi masih jauh di bawah jumlah kasus pada dua desa dalam cluster tinggi.

Cluster tinggi terdiri atas Limbangan dan Kradenan. Kedua desa tersebut memiliki jumlah kasus paling besar, yaitu masing-masing 79 dan 76 kasus. Tingginya jumlah kasus pada

kedua wilayah tersebut menyebabkan centroid cluster tinggi mencapai 77,50 kasus.

Dari perspektif pengelolaan kesehatan wilayah, hasil clustering dapat dimanfaatkan sebagai informasi awal dalam menentukan prioritas pemantauan. Desa dalam cluster tinggi dapat menjadi prioritas dalam kegiatan pemantauan dan evaluasi program pengendalian Diabetes Mellitus. Sementara itu, desa dalam cluster sedang dapat menjadi wilayah yang perlu mendapatkan perhatian preventif agar peningkatan kasus dapat dikendalikan.

Adapun desa dalam cluster rendah tetap perlu mendapatkan pemantauan karena kategori rendah tidak berarti bahwa wilayah tersebut bebas dari risiko Diabetes Mellitus. Kategori tersebut hanya menunjukkan bahwa jumlah kasus yang tercatat lebih rendah dibandingkan dengan wilayah lain dalam dataset penelitian.

Hasil penelitian juga memperlihatkan adanya ketidakseimbangan jumlah kasus antarwilayah. Dua desa yang masuk cluster tinggi memiliki 155 kasus atau 42,47% dari total 365 kasus. Kondisi ini menunjukkan bahwa analisis pengelompokan dapat memberikan informasi yang berbeda dibandingkan hanya melihat rata-rata keseluruhan wilayah.

Rata-rata keseluruhan jumlah kasus per desa adalah:

$$\frac{365}{13} = 28,08$$

Namun, rata-rata tersebut tidak sepenuhnya menggambarkan kondisi masing-masing desa karena terdapat perbedaan jumlah kasus yang cukup besar. Desa Sindangjaya memiliki 8 kasus, sedangkan Limbangan memiliki 79 kasus. Oleh karena itu, pengelompokan berbasis K-Means dapat memberikan gambaran yang lebih terstruktur mengenai variasi tingkat kasus antarwilayah.

Dalam pengembangan penelitian lebih lanjut, hasil clustering dapat diintegrasikan dengan Sistem Informasi Geografis atau Web GIS. Setiap desa dapat ditampilkan dalam bentuk peta tematik berdasarkan cluster yang diperoleh. Dengan demikian, pengguna dapat melihat distribusi spasial wilayah rendah, sedang, dan tinggi secara lebih intuitif.

5. KESIMPULAN

Berdasarkan penelitian mengenai penerapan metode K-Means pada pengelompokan persebaran kasus Diabetes Mellitus di Kecamatan Kersana, dapat disimpulkan bahwa metode K-Means dapat digunakan untuk mengelompokkan 13 desa berdasarkan jumlah kasus Diabetes Mellitus ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi.

Hasil pengelompokan menghasilkan 9 desa dalam cluster rendah, yaitu Cigedog, Kubangpari, Ciampel, Jagapura, Sutamaja, Kemukten, Sindangjaya, Pande, dan Keramatsampang. Cluster sedang terdiri atas Cikandang dan Kersana. Sementara itu, cluster tinggi terdiri atas Limbangan dan Kradenan.

Nilai centroid akhir masing-masing kelompok adalah 15,33 kasus untuk cluster rendah, 36 kasus untuk cluster sedang, dan 77,50 kasus untuk cluster tinggi. Hasil tersebut menunjukkan adanya perbedaan tingkat kasus yang cukup jelas antarwilayah.

Desa Limbangan dan Kradenan menjadi wilayah dengan tingkat kasus tertinggi. Kedua desa tersebut memiliki total 155 kasus atau sekitar 42,47% dari keseluruhan 365 kasus. Kondisi tersebut menunjukkan bahwa kedua wilayah dapat menjadi prioritas dalam kegiatan pemantauan dan perencanaan program pengendalian Diabetes Mellitus.

Evaluasi menggunakan *Silhouette Score* menghasilkan nilai 0,690 yang menunjukkan bahwa struktur cluster yang

dihasilkan cukup baik. Dengan demikian, metode K-Means dapat digunakan sebagai pendekatan analisis untuk mengelompokkan wilayah berdasarkan tingkat jumlah kasus Diabetes Mellitus.

Hasil penelitian ini dapat dikembangkan lebih lanjut dengan mengintegrasikan hasil clustering ke dalam Web GIS sehingga informasi mengenai wilayah dengan tingkat kasus rendah, sedang, dan tinggi dapat divisualisasikan dalam bentuk peta interaktif.

6. SARAN

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan adalah sebagai berikut:

1. Penelitian selanjutnya dapat menggunakan data kasus Diabetes Mellitus dari beberapa tahun sehingga dapat dilakukan analisis perubahan cluster dari waktu ke waktu.
2. Variabel penelitian dapat dikembangkan dengan menambahkan jumlah penduduk, prevalensi, usia penderita, jenis kelamin, serta faktor risiko lainnya.
3. Pengelompokan berdasarkan jumlah kasus dapat dikembangkan menjadi pengelompokan berdasarkan rasio atau prevalensi terhadap jumlah penduduk sehingga perbandingan antarwilayah menjadi lebih proporsional.
4. Hasil clustering dapat diintegrasikan dengan Web GIS untuk memberikan visualisasi spasial mengenai persebaran kasus Diabetes Mellitus.
5. Penelitian selanjutnya dapat membandingkan K-Means dengan metode clustering lainnya untuk mengetahui metode yang paling sesuai dengan karakteristik data.

DAFTAR PUSTAKA

- Aminuddin, A., Sima, Y., Izza, N. C., Norma, N. S., & Arda, D. (2023). *Abdimas Polsaka : Jurnal Pengabdian Kepada Masyarakat Edukasi Kesehatan Tentang Penyakit Diabetes Melitus bagi Masyarakat*. 7–12.
- Irawan, D., Wijaya, G., & Warisaji, T. T. (2025). *Penerapan Algoritma K-Means Clustering untuk Segmentasi Nasabah Bank*. 6(1).
- Ishak, R. (2024). *Optimasi K-Means pada Clustering Penyakit Ibu Hamil Menggunakan Random Forest Optimization of K-Means in Disease Clustering of Pregnant Women Using Random Forest*. 7, 41–47.
- Korespondensi, P. (2025). *Departemen Keperawatan , Fakultas Keperawatan dan Kebidanan , Universitas Departemen Kedokteran , Fakultas Kedokteran , Universitas Nahdlatul Ulama Surabaya , Prodi Pendidikan Profesi Ners , STIKES Adi Husada Surabaya , Surabaya Indonesia , termasuk di Kelurahan Wonokromo , Surabaya . Penyakit ini berkaitan erat besar penderita diabetes di Indonesia . Di Kelurahan Wonokromo , terdapat posyandu lansia yang belum banyak di manfaatkan oleh masyarakat untuk melakukan pemeriksaan kesehatan terutama pemeriksaan kadar glukosa darah . Hasil penelitian menunjukan bahwa sebagian besar (63 , 0 %) penderita diabetes mellitus di kelurahan wonokromo memiliki pola kesehatan yang cukup serius . Tingginya prevalensi diabetes di wilayah ini dipengaruhi oleh Banyak penderita diabetes yang belum disiplin dalam menjaga pola makan , masih mengonsumsi makanan tinggi gula dan karbohidrat , serta kurang melakukan aktivitas fisik yang dapat*

membantu mengontrol kadar gula darah . Selain itu , ketidakpatuhan dalam utama dalam pengelolaan diabetes . Keterbatasan akses terhadap layanan kesehatan yang memadai , serta kurangnya pengetahuan mengenai pentingnya monitoring gula darah secara rutin , turut memperburuk kondisi ini . Semua faktor tersebut menjadi kendala besar. 3, 92–104.

- Luhur, U. B., Raya, J. C., Utara, P., & Selatan, J. (2020). *Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat*. 1,2. 7(2).
- Mafela, G., Sujak, M., Rofiq, H. N., & Tawakal, F. I. (2025). *Implementation of K-Means Clustering for Optimizing Non-Communicable Disease Budgets Implementasi K-Means Clustering untuk Optimalisasi Anggaran Penyakit Tidak Menular*. 5(January), 67–74.
- Magelang, U. M. (2025). *Universitas Muhammadiyah Magelang*.
- Norshahlan, M., Jaya, H., & Kustini, R. (2023). *Penerapan Metode Clustering Dengan Algoritma K-means Pada Pengelompokan Data Calon Siswa Baru*. 2(November), 1042–1053.
- Studi, P., Informasi, S., Komputer, F. I., Pamulang, U., & Selatan, T. (2026). *Penerapan Teknik Clustering Menggunakan Algoritma K-Means Berbasis RFM untuk Segmentasi Pelanggan pada Dataset Online Retail II*. 02(03), 668–673.
- Suryani, N., Priyanti, E., & Sihombing, G. (2026). *Implementasi Algoritma K-Means Clustering dalam Pengelompokan Potensi Lokasi Usaha Waralaba Berdasarkan Karakteristik Wilayah*. 7(2).