

# **PENGELOMPOKAN WILAYAH PENYEBARAN COVID-19 DI KABUPATEN KARAWANG MENGGUNAKAN ALGORITMA K-MEANS**

*Zulia Imami Alfianti  
Universitas Bina Sarana Informatika Kampus Kabupaten Karawang  
Jl. Banten No. 1, Karangpawitan, Karawang 41371, Jawa Barat  
zulia.zim@bsi.ac.id*

## **Abstrak**

*Virus Covid-19 merupakan keluarga besar coronavirus yang dapat menyerang hewan. Covid-19 merupakan penyakit menular yang ditandai oleh gejala pada bagian pernapasan akut (Coronavirus 2 (Severe Acute respiratory Syndrome Coronavirus 2) atau SARS-CoV-2). Virus ini menyerang penduduk di Wuhan China sejak Desember 2019. Penularan virus ini terjadi jika adanya kontak langsung antar manusia. Sejak Februari 2020 penyebaran virus covid-19 mulai terjadi di berbagai wilayah di Indonesia dan semakin meningkat setiap harinya. Pada penelitian ini akan dilakukan pengklasteran wilayah penyebaran Covid-19 di Kabupaten Karawang. Pengklasteran dilakukan bertujuan untuk mengelompokkan himpunan data yang besar untuk memudahkan analisis data atau pengolahan data lebih lanjut menggunakan metode k-means dengan membagi data kecamatan di kabupaten Karawang menjadi tiga cluster yaitu cluster wilayah penyebaran rendah, cluster wilayah penyebaran sedang, dan wilayah penyebaran tinggi. Hasil dari penelitian ini yaitu terdapat 50% persen wilayah yang tingkat penyebarannya rendah, 33,3% persen wilayah yang tingkat penyebarannya sedang, dan 16,7% persen wilayah yang tingkat penyebarannya tinggi.*

**Kata Kunci:** *Clustering, Covid-19, Data Mining, K-Means*

## **Abstract**

*The Covid-19 virus is a big family of coronaviruses that can infect animals. Covid-19 is an infectious disease characterized by symptoms in the acute respiratory tract (Coronavirus 2 (Severe Acute Respiratory Syndrome Coronavirus 2) or SARS-CoV-2). This virus has attacked residents in Wuhan, China since December 2019 The transmission of this virus occurs if there is direct contact between humans. Since February 2020 the spread of the Covid-19 virus has begun to occur in various regions in Indonesia and is increasing every day. In this study, clustering of the Covid-19 distribution area in Karawang Regency will be carried out. Clustering is carried out with the aim of grouping big data sets to facilitate data analysis or further data processing using the k-means method by dividing sub-district data in Karawang district into three clusters, namely clusters of low distribution areas, clusters of distribution areas, medium, and high distribution areas. The results of this study are that there are 50% percent of areas with low distribution, 33.3% of areas with moderate distribution, and 16.7% percent of areas with high distribution.*

**Keywords:** *Clustering, Covid-19, Data Mining, K-Means*

## PENDAHULUAN

Virus Covid-19 merupakan penyakit menular yang ditandai oleh gejala pada bagian pernapasan akut (*Coronavirus 2 (Severe Acute Respiratory Syndrome Coronavirus 2)*) atau SARS-CoV-2) [1]. Infeksi Virus Corona atau Covid-19 disebabkan oleh Corona virus, yaitu kelompok virus yang menginfeksi sistem pernapasan, pada sebagian besar kasus corona virus hanya menyebabkan infeksi pernapasan ringan sampai sedang, seperti flu, akan tetapi, virus ini juga bisa menyebabkan infeksi pernapasan berat, seperti Pneumonia, *MiddleEast Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome* (SARS) [2]. Virus ini dapat menyerang hewan dan juga manusia dan pada manusia gejalanya berupa infeksi yang serupa dengan penyakit SARS dan MERS, hanya saja Covid-19 bersifat lebih masif perkembangannya. Saat ini, penyebaran SARS-CoV-2 dari manusia ke manusia menjadi sumber transmisi utama sehingga penyebaran menjadi lebih agresif. Transmisi SARS-CoV-2 dari pasien simptomatik terjadi melalui droplet yang keluar saat batuk atau bersin [3]. Indonesia juga merupakan salah satu negara yang terdampak wabah yang satu ini [4]. Keberadaan virus Covid-19 ini tentunya melumpuhkan semua kegiatan baik kegiatan sosial dan maupun kegiatan perekonomian [5]. Penyebaran virus Covid-19 di seluruh

provinsi di Indonesia menyebabkan berbagai dampak pada semua bidang yang tentunya menghambat aktivitas manusia. Beberapa wilayah ditetapkan sebagai zona merah yang artinya wilayah yang memiliki tingkat penderita dan penyebarannya sangat tinggi. Sempat beberapa kali dikategorikan sebagai wilayah zona merah yang artinya wilayah yang tingkat penderitanya tinggi, membuat wilayah Kabupaten Karawang harus mendapat penanganan lebih cepat oleh pemerintah daerah Kabupaten Karawang [6]

Penelitian ini akan mengelompokkan wilayah penyebaran virus Covid-19 di Kabupaten Karawang. *Data mining* yang biasa disebut sebagai *Knowledge Discovery in Database* (KDD) merupakan kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola hubungan dalam himpunan data yang berukuran besar [7]. Hasil perhitungan dari *data mining* ini dapat digunakan untuk pengambilan keputusan. Teknik dalam *data mining* yang digunakan untuk melakukan pengelompokkan sejumlah data atau objek ke dalam *cluster* (grup) disebut dengan *clustering*. *Clustering* ini memungkinkan melakukan pengelompokkan data wilayah penyebaran Covid-19 menjadi menjadi beberapa *cluster* (grup).

Penelitian serupa mengenai metode *clustering* telah dilakukan oleh beberapa peneliti diantaranya Nayuni Diwitri dkk. [8] mengenai penerapan algoritma *K-Means*

dalam menentukan tingkat penyebaran Covid-19 di Indonesia. Hasil penelitian tersebut yaitu dari 34 provinsi di Indonesia terbagi menjadi pusat *cluster* yang diperoleh yaitu untuk *cluster* pertama atau penyebaran tingkat tinggi berada pada Provinsi DKI Jakarta, pusat *cluster* kedua atau penyebaran tingkat sedang berada pada provinsi Jawa Barat, dan pusat klaster ketiga atau penyebaran tingkat sedang berada pada Provinsi Banten. Penelitian selanjutnya oleh M. Hasnil Aditya dkk. [9] mengenai Penggunaan algoritma *K-Means* untuk *clustering* obat-obatan di RSUD Pekanbaru. Dari hasil klasterisasi pada obat-obatan tersebut kelompok obat yang termasuk pemakaian sedikit rata-rata permintaan obat setiap tahunnya kurang dari 18000 buah, dan obat yang termasuk pemakaian sedang rata-rata permintaan obat setiap tahunnya diantara 18000–70000 buah, sedangkan obat yang masuk kedalam kelompok obat yang pemakaian tinggi rata-rata permintaan obat setiap tahunnya diatas 70000 buah. Penelitian yang ketiga oleh Muhammad Rafi Muttaqin dan Meriska Defriani [10] mengenai penggunaan algoritma *K-Means* untuk pengelompokan topik skripsi mahasiswa hasil pengelompokan dengan nilai *cluster* yang paling tinggi dapat menunjukkan kemampuan mahasiswa pada tiap kelompok bidang keahlian. Hasil pengelompokan dianalisis dengan melihat jumlah matakuliah pada setiap kelompok bidang keahlian. Jumlah matakuliah yang paling banyak pada suatu kelompok bidang keahlian menandakan

bahwa mahasiswa memiliki kemampuan lebih baik pada bidang tersebut sehingga direkomendasikan topik skripsi yang sesuai dengan kelompok bidang keahlian tersebut.

Penelitian ini bertujuan untuk menentukan pengelompokan wilayah berdasarkan tingkat penyebaran Covid-19 di wilayah Kabupaten Karawang menggunakan metode *clustering* algoritma *K-Means*.

## METODE PENELITIAN

Penelitian dilakukan dengan menggunakan metode klasifikasi *K-Means*. Metode klasifikasi *K-Means* merupakan algoritma yang melakukan pengelompokan data berdasarkan data yang terdekat dengan titik pusat *cluster*.

Data yang digunakan dalam penelitian ini adalah data penyebaran virus Covid-19 di kabupaten Karawang yang diperoleh dari laman covid19.karawangkab.go.id. data yang digunakan adalah data penyebaran yang direkap pada tanggal 20 Mei 2021.

*Data mining* disebut juga *Knowledge Discovery in Database* (KDD) melakukan ekstraksi informasi dari tumpukan data. Proses pencarian informasi tersebut akan menemukan pola kecenderungan dari data yang kemudian hasil dari penambangan (*mining*) dapat menjadi informasi yang mudah dipahami.

Analisis *cluster* merupakan salah satu teknik *data mining* yang bertujuan untuk mengidentifikasi sekelompok objek yang

mempunyai kemiripan karakteristik tertentu yang dapat dipisahkan dengan kelompok objek lainnya, sehingga objek yang berada dalam kelompok yang sama relatif lebih homogen daripada objek yang berada pada kelompok yang berbeda [11].

*K-Means* yaitu salah satu dari metode pengelompokan data nonhierarki (sekatan) yang dapat mempartisi data kedalam bentuk dua kelompok ataupun lebih [9]. *K-Means clustering* merupakan sebuah konstanta dari sejumlah *cluster* yang diinginkan, sedangkan *Means* atau dapat didefinisikan sebagai *cluster* adalah suatu nilai rata-rata dari sekumpulan populasi data [12]. Metode *K-Means* digunakan untuk mengelompokkan data ke dalam beberapa kelompok, data dalam satu kelompok mempunyai karakteristik yang sama dan mempunyai karakteristik yang berbeda dengan data dari kelompok yang lain [13]. Metode ini berusaha meminimalkan variasi antar data yang ada di dalam suatu *cluster* dan memaksimalkan variasi data yang ada di *cluster* lainnya.

Langkah-langkah *clustering* menggunakan algoritma *K-Means* [14]:

1. Tentukan Jumlah *cluster*
2. Alokasikan data ke dalam *cluster* secara acak
3. Menghitung centroid/rata-rata dari data yang ada di masing-masing *cluster*
4. Alokasikan masing-masing data ke *centroid/cluster* terdekat
5. Jika masih ada data yang berpindah *cluster* adanya selisih antara nilai *centroid* lama

dengan *centroid* baru maka kembali lakukan perhitungan *centroid* kembali di masing-masing *cluster*. Jika tidak ada perubahan terhadap nilai *centroid* maka perhitungan dihentikan.

Proses mengolah data yang telah didapatkan dengan menyusun pola serta mengelompokkan pola tersebut menjadi beberapa *cluster* dengan metode *K-means*. Pengelompokkan pola menjadi beberapa *cluster* menggunakan rumus *euclidian distance*. Dengan menghitung jarak antara data dengan *centroid* sama dengan akar dari hasil penjumlahan antara data pada atribut ke-*n* dikurangi objek data pada *centroid* ke-*k*. Hasil jarak ini yang akan digunakan untuk mencari data mana yang memiliki tingkat kemiripan berdasarkan jarak yang paling kecil/terpendek. Perhitungan menggunakan rumus *euclidian distance* diformulasikan dengan menggunakan persamaan 1.

$$d(a, b) = \sqrt{\sum_{k=1}^n (a_k - b_k)^2} \quad (1)$$

Dimana:

d = jarak

n = jumlah dimensi (atribut)

$a_k$  dan  $b_k$  = atribut ke- *k* dari objek data

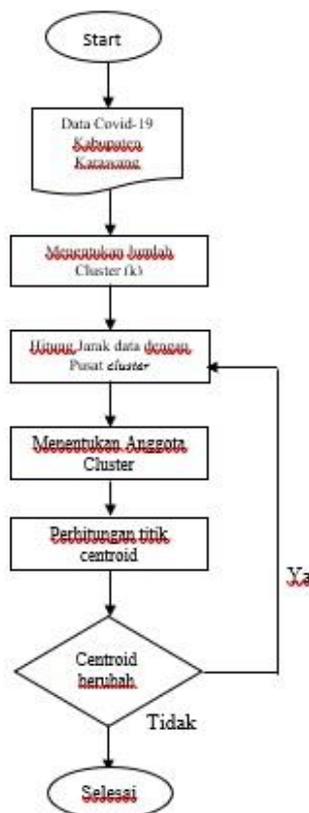
Untuk melengkapi pengetahuan dasar dalam melakukan penelitian ini digunakan studi literatur yang bersumber dari jurnal maupun penelitian terkait.

## HASIL DAN PEMBAHASAN

Proses pengelompokan daerah terdampak Covid-19 di Kabupaten Karawang menggunakan algoritma *K-Means*. Langkah pertama yang dilakukan adalah menentukan jumlah *cluster* yang ditentukan yaitu jumlah *cluster* harus lebih sedikit dari jumlah data yang akan diteliti. Selanjutnya, pilih titik *centroid* secara acak yang nantinya titik ini akan menjadi *centroid* dari masing-masing *cluster*. Kemudian hitung jarak dan alokasikan masing-masing data ke *centroid* terdekat. Setelah itu, tentukan *centroid* baru dari data yang ada di masing-masing *cluster*. Apabila masih ada data yang berpindah *cluster* atau ada perubahan nilai *centroid*, maka kembali hitung jarak dan alokasikan kembali data ke masing-masing *centroid*. Jika tidak ada perpindahan *cluster* atau perubahan

nilai *centroid* maka hentikan proses *clustering*. Langkah-langkah dalam metode *K-Means* dari data yang akan diteliti digambarkan dalam *flowchart* yang dapat dilihat pada Gambar 1.

Langkah pertama yang dilakukan pada proses *clustering* menggunakan metode *K-Means* adalah menentukan jumlah *cluster* awal secara bebas yaitu 3 *cluster* dan *centroid* dari data 30 Kecamatan di Kabupaten Karawang yang ditandai dengan K1 sampai dengan K30. Ditentukan 3 titik *centroid* awal yaitu K17, K18, K27 diambil dari jumlah kasus tinggi, sedang/menengah, dan rendah. Data kecamatan terdampak Covid-19 di Kabupaten Karawang dapat dilihat pada Tabel 1.



**Gambar 1. Langkah-Langkah Pengklusteran Menggunakan Metode *K-Means***

**Tabel 1. Data Kecamatan Terdampak Covid-19 Kabupaten Karawang pada Tanggal 20 Mei 2021**

Kode	Kecamatan	Perawatan	Isolasi	Sembuh	Meninggal
K1	Banyusari	1	0	165	7
K2	Batujaya	1	0	112	8
K3	Ciampel	2	0	307	6
K4	Cibuaya	1	0	154	9
K5	Cikampek	13	0	1016	41
K6	Cilamaya Kulon	0	0	137	7
K7	Cilamaya Wetan	3	0	380	18
K8	Cilebar	1	0	112	2
K9	Jatisari	3	0	403	18
K10	Jayakarta	0	0	178	13
K11	Karawang Barat	12	4	2378	62
K12	Karawang Timur	8	0	2100	54
K13	Klari	8	7	1910	68
K14	Kota Baru	6	0	1390	37
K15	Kutawaluya	1	0	144	10
K16	Lemahaban g	1	0	217	23
K17	Majalaya	3	0	990	20
K18	Pakisjaya	0	0	19	6
K19	Pangkalan	0	0	82	2
K20	Pedes	0	0	172	14
K21	Purwasari	6	0	750	21
K22	Rawamerta	1	0	314	18
K23	Rengasdeng klok	2	4	462	28
K24	Tegalwaru	0	0	134	3
K25	Telagasari	1	0	415	32
K26	Telukjambe Barat	5	5	345	9
K27	Telukjambe Timur	8	1	3246	73
K28	Tempuran	2	0	169	9
K29	Tirtajaya	0	0	147	6
K30	Tirtamulya	0	0	102	4

Sumber : covid19.karawangkab.go.id

Dari data penyebaran Covid-19 30 Kecamatan di Kabupaten Karawang tersebut ditentukan tiga *cluster* sebagai titik *centroid* awal yaitu data ke-27, 17, dan 18 dapat dilihat pada Tabel 1. dipilih berdasarkan jumlah kasus keseluruhan dengan menjumlahkan atribut dari masing-masing kecamatan

sehingga didapatkan satu kecamatan yang memiliki kasus tertinggi, satu kecamatan yang memiliki kasus sedang/menengah, dan satu kecamatan yang memiliki kasus terendah dari 30 kecamatan di Kabupaten Karawang yang ditetapkan sebagai *centroid* awal. *Centroid* awal yang telah ditentukan dapat dilihat pada

Tabel 2. Setelah dilakukan penentuan *centroid* awal, maka langkah selanjutnya adalah menghitung jarak antara *centroid* awal dengan titik *centroid* menggunakan rumus sebagai berikut. Berdasarkan persamaan (1). Hasil yang didapatkan untuk menentukan jarak antara *centroid* awal dengan titik *centroid* menggunakan rumus *euclidian*

*distance* yang hasilnya terdapat pada Tabel 3 ini, kemudian dipilih *centroid* mana dari masing-masing kode yang memiliki jarak terpendek.

$$d_{(a,b)} = \sqrt{(1-5)^2 + (0-5)^2 + (165-345)^2 + (7-9)^2}$$

$$d_{(a,b)} = \sqrt{(1-5)^2 + (0-5)^2 + (165-345)^2 + (7-9)^2}$$

$$d_{(a,b)} = \sqrt{(1-0)^2 + (0-0)^2 + (165-19)^2 + (7-6)^2}$$

**Tabel 2. Hasil Pemilihan Centorid Awal**

Centorid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	5	5	345	9
C2	3	0	990	20
C3	0	0	19	6

**Tabel 3. Hasil Perhitungan Jarak Cluster dengan Titik Centroid**

Kode	C1	C2	C3	Kode	C1	C2	C3
K1	180	825	146	K16	128.9	773	198.7
K2	233	878	93	K17	645	0	971
K3	38.5	683	288	K18	326	971	0
K4	191	836	135	K19	263	908	63
K5	671.8	34.9	997.7	K20	173	818	153
K6	208	853	118	K21	405	240	731
K7	36.5	610	361	K22	32.9	676	295
K8	233	878	93	K23	118.5	528	443.5
K9	58.9	587	384	K24	211	856	115
K10	167	812	159	K25	73.9	575	396.8
K11	2033.7	1388.6	2359.6	K26	0	645	326
K12	1755.5	1110.5	2081.5	K27	2901.7	2256.6	3227.7
K13	1566	921	1892	K28	176	821	150
K14	1045	400	1371	K29	198	843	128
K15	201	846	125	K30	243	888	83

**Tabel 4. Keanggotaan Cluster Hasil Perhitungan Centroid pada Iterasi 1**

Centorid	Keanggotaan Cluster													
C1	K3	K7	K9	K16	K22	K23	K25	K26						
C2	K5	K11	K12	K13	K14	K17	K21	K27						
C3	K1	K2	K4	K6	K8	K10	K15	K18	K19	K20	K24	K28	K29	K30

**Tabel 5. Nilai Centroid dari Hasil Perhitungan Iterasi 1**

Centorid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	2	1.125	355.375	19
C2	8	1.75	1722.5	47
C3	0.5	0	131	7.1

Dari hasil perhitungan untuk mencari jarak terpendek pada Tabel 3 didapatkan jarak terpendek antara data dengan *centroid* sehingga didapatkan keanggotaan *cluster* yaitu C1 sebanyak 8, C2 sebanyak 8, dan C3 sebanyak 14, dapat dilihat pada Tabel 4. Langkah selanjutnya yaitu dengan kembali menghitung nilai rata-rata anggota *cluster* sehingga didapatkan nilai *centroid* iterasi 1 yang dapat dilihat pada Tabel 5: Hasil yang didapatkan di iterasi 1 ternyata ada perubahan nilai *centroid*, maka diharuskan melakukan perhitungan kembali menggunakan rumus *eucledian distance* untuk menentukan jarak antara *centroid* dengan titik *centroid* yang

hasilnya terdapat pada Tabel 6, kemudian dipilih kembali *centroid* mana dari masing-masing kode yang memiliki jarak terpendek, hasil perhitungan tersebut dapat dilihat pada Tabel 6.

Berdasarkan hasil perhitungan nilai *centroid* pada iterasi 2 terdapat perbedaan antara nilai *centroid* yang dihasilkan pada iterasi 1. Sehingga diperlukan iterasi berikutnya dengan melakukan perhitungan kembali jarak antara data dengan nilai *centroid* yang terakhir untuk kemudian ditentukan kembali keanggotaan *cluster* melalui jarak terkecil antara data dengan *cluster* serta hitung nilai *centroid* berikutnya.

**Tabel 6. Hasil Perhitungan Jarak Cluster dengan Titik Centroid Sebelumnya Untuk Menentukan Centroid pada Iterasi 2**

Kode	C1	C2	C3	Kode	C1	C2	C3
K1	190.7	1558	34.5	K16	138.4	1505.7	87.9
K2	243	1610.9	18.5	K17	644.6	733	859.5
K3	50	1416	176.5	K18	336.6	1704	111.5
K4	201.6	1568.9	23.5	K19	273.9	1641	48.7
K5	661	706.5	886	K20	183.4	1550.8	42
K6	218.7	1586	6.5	K21	394.6	972.8	619.6
K7	24.6	1342.8	249.7	K22	41	1408.8	183.8
K8	243.9	1611	19	K23	107	1260.6	332
K9	47.6	1342.8	272.7	K24	221	1589	5
K10	177	1544.8	47.8	K25	61	1307.6	285.5
K11	2023	655.6	2248	K26	15	1378	214.6
K12	1744.9	377.5	1970	K27	2891	1523.7	3116
K13	1555	188.7	1780	K28	186.6	1553.9	38.5
K14	1034.7	332.6	1259.8	K29	208.7	1576	16.5
K15	211.5	1578.9	13.8	K30	253.8	1621	28.6

**Tabel 7. Nilai Centroid yang Dihasilkan pada Iterasi 2**

Centorid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	4	1.1	538.2	21.4
C2	8.4	2.4	2204.8	58.8
C3	0.53	0	8	8.2

Berikut ini nilai *centroid* dari iterasi 2 yang dapat dilihat pada Tabel 7. Hasil yang didapatkan untuk menentukan jarak antara *centroid* dengan titik *centroid* menggunakan rumus *euclidian distance* pada iterasi 3 yang hasilnya dijabarkan pada Tabel 8, kemudian dipilih *centroid* mana dari masing-masing kode yang memiliki jarak terpendek, hasil perhitungan tersebut yang dapat dilihat pada Tabel 8. Setelah dilakukan perhitungan jarak data dengan titik *centroid*, didapatkan hasil perhitungan *centroid* pada iterasi 3. Berikut ini nilai *centroid* dari iterasi 3 yang dapat

dilihat pada Tabel 9. Setelah dilakukan perhitungan, proses iterasi terhenti di iterasi 3 karena nilai *centroid* iterasi 3 sama dengan nilai *centroid* iterasi 2. Perbandingan nilai *centroid* yang berasal dari iterasi awal sampai dengan iterasi 3 tersebut dapat dilihat pada Tabel 10.

Setelah dilakukan perbandingan nilai *centroid* pada Tabel 10, proses iterasi berhenti pada iterasi 3 karena nilai *centroid* pada iterasi 2 sama dengan nilai *centroid* pada iterasi 3. Sehingga didapatkan keanggotaan *cluster* seperti ditampilkan pada Tabel 11.

**Tabel 8. Hasil Perhitungan Jarak Cluster dengan Titik Centroid Sebelumnya Untuk Menentukan Centroid pada Iterasi 3**

Kode	C1	C2	C3	Kode	C1	C2	C3
K1	373	2040	156.8	K16	321	1988	209
K2	426	2093	103.8	K17	451.8	1215	981.8
K3	231.7	1898.5	298.8	K18	519	2186	11
K4	384	2051	145.8	K19	456.6	2123	74
K5	478	1188.9	1008	K20	366	2033	163.9
K6	401	2068	128.8	K21	211.8	1455	741.9
K7	158.2	1825	371.9	K22	224	1891	305.9
K8	426.6	2093.5	103.9	K23	76.5	1743	454
K9	135	1802	394.9	K24	404.6	2071.5	125.9
K10	360	2027	169.8	K25	123.6	1790	407
K11	1840	173	2370	K26	193.6	1860	336.8
K12	1562	104.9	2092	K27	2708	1041	3238
K13	1372.6	294.9	1902.7	K28	369	2036	160.8
K14	851.9	851	1382	K29	391.5	2058	138.8
K15	394	2061	135.8	K30	436.5	2103.5	93.8

**Tabel 9. Nilai Centroid yang Dihasilkan pada Iterasi 3**

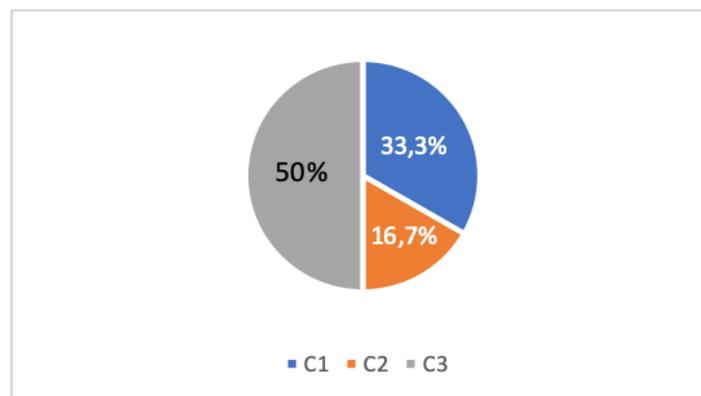
Centorid Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4
C1	4	1.1	538.2	21.4
C2	8.4	2.4	2204.8	58.8
C3	0.53	0	8	8.2

**Tabel 10. Perbandingan Nilai *Centroid* pada Iterasi Awal sampai Iterasi 3**

Centorid		Keanggotaan Cluster			
Awal	Atribut 1	Atribut 2	Atribut 3	Atribut 4	
Iterasi 0					
C1	5	5	345	9	
C2	5	0	990	20	
C3	0	0	19	6	
Iterasi 1					
C1	2	1.125	355.375	19	
C2	8	1.75	1722.5	47	
C3	0.5	0	131	71	
Iterasi 2					
C1	4	1.1	538.2	21.4	
C2	8.4	2.4	2204.8	58.8	
C3	0.53	0	8	8.2	
Iterasi 3					
C1	4	1.1	538.2	21.4	
C2	8.4	2.4	2204.8	58.8	
C3	0.53	0	8	8.2	

**Tabel 11. Keanggotaan *Cluster* Hasil Perhitungan *Centroid* pada Iterasi 3**

Centorid	Keanggotaan Cluster														
C1	K3	K5	K7	K9	K17	K21	K22	K23	K25	K26					
C2	K11	K12	K13	K14	K27										
C3	K1	K2	K4	K6	K8	K10	K15	K16	K18	K19	K20	K24	K28	K29	K30



**Gambar 2. Diagram Hasil Pengklasteran Wilayah Terdampak Covid-19 Kabupaten Karawang**

Dari hasil *clustering* wilayah penyebaran Covid-19 diatas, maka untuk lebih jelasnya penulis gambarkan pada diagram lingkaran yang ditampilkan pada Gambar 2. Hasil perhitungan menggunakan metode *K-Means* pada diagram lingkaran yang terdapat pada Gambar 2, dari 30

kecamatan di wilayah Kabupaten Karawang yang telah melewati proses *clustering* menggunakan metode *K-Means* dibagi menjadi tiga *cluster* penyebaran yang dikategorikan menjadi tiga wilayah tingkat penyebaran tinggi yang memiliki tingkat kasus tertinggi, wilayah yang memiliki

tingkat penyebaran sedang/menengah yang memiliki jumlah kasus diantara tinggi dan rendah, dan wilayah yang memiliki tingkat wilayah penyebaran rendah memiliki tingkat kasus terendah.

Pada diagram lingkaran diatas yang dihasilkan dari proses *clustering* terdapat 16,7% yaitu 5 kelurahan yang memiliki tingkat penyebaran tinggi, 33,3% terdiri dari 10 kelurahan memiliki tingkat penyebaran sedang/menengah karena berada diantara persentasi tingkat tinggi dan rendah, dan 50% terdiri dari 15 kelurahan memiliki tingkat penyebaran rendah.

## KESIMPULAN DAN SARAN

Setelah dilakukan proses pengelompokan wilayah terdampak Covid-19 di Kabupaten Karawang dengan menggunakan metode *K-Means* dapat disimpulkan bahwa dari 30 kecamatan yang ada di Kabupaten Karawang 50% wilayah termasuk kedalam kategori terdampak rendah, 33,3% wilayah termasuk kedalam kategori terdampak sedang, dan 16,7% wilayah termasuk kedalam kategori terdampak tinggi.

Mengingat pandemi Covid-19 yang masih berlangsung maka diperlukan pengkajian lanjutan mengenai wilayah terdampak Covid-19 di Kabupaten Karawang dengan menggunakan data-data yang telah diperbaharui setiap harinya berdasarkan kasus yang terjadi di masing-masing kecamatan Kabupaten Karawang.

Untuk mengembangkan penelitian ini, dibutuhkan penelitian lanjutan metode *clustering* terhadap wilayah penyebaran Covid-19 yang lebih luas yaitu pada 34 provinsi di Indonesia yang dilakukan secara lebih *valid* dengan ditetapkan nilai *centroid* terbaik untuk mengetahui pengelompokan wilayah berdasarkan tingkat penyebaran Covid-19 di Indonesia.

## DAFTAR PUSTAKA

- [1] C. C. Lai, T. P. Shih, W. C. Ko, H. J. Tang, and P. R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges," *Int. J. Antimicrob. Agents*, vol. 55, no. 3, p. 105924, 2020, doi: 10.1016/j.ijantimicag.2020.105924.
- [2] D. Indriyani and C. Virus, "Penanganan dan Pencegahan Pandemi Wabah Virus Corona ( Covid-19 )," vol. 2, 2020.
- [3] A. Susilo *et al.*, "Coronavirus Disease 2019: Tinjauan Literatur Terkini," *J. Penyakit Dalam Indones.*, vol. 7, no. 1, p. 45, 2020, doi: 10.7454/jpdi.v7i1.415.
- [4] I. Wahidah, M. A. Septiadi, M. C. A. Rafqie, N. Fitria, S. Hartono, and R. Athallah, "Pandemik Covid-19 : Analisis Perencanaan Pemerintah dan Masyarakat dalam Berbagai Upaya Pencegahan COVID-19 Pandemic :

- Analysis of Government and Community Planning in Various Prevention Measures,” vol. 11, no. 3, pp. 179–188, 2020.
- [5] C. Fikri, “Tiga Dampak Pandemi Covid-19 Bagi Perekonomian Nasional,” *Beritasatu.com*, p. 2021, 2021, [Online]. Available: <https://www.beritasatu.com/ekonomi/728997/tiga-dampak-pandemi-covid19-bagi-perekonomian-nasional>.
- [6] Y. F. Silitonga, “Seluruh Kecamatan di Karawang Zona Merah,” p. 2021, 2021.
- [7] S. Handoko, F. Fauziah, and E. T. E. Handayani, “Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering,” *J. Ilm. Teknol. dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020, doi: 10.35760/tr.2020.v25i1.2677.
- [8] N. Dwitri, J. A. Tampubolon, S. Prayoga, F. Ilmi Zer, and D. Hartama, “Penerapan Algoritma K-Means Dalam Menentukan Tingkat Kepuasan Pembelajaran Online Pada Masa Pandemi Covid-19 di Indonesia,” *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 101–105, 2020.
- [9] M. H. Adiya and Y. Desnelita, “Jurnal Nasional Teknologi dan Sistem Informasi Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru,” vol. 01, pp. 17–24, 2019.
- [10] M. R. Muttaqin and M. Defriani, “Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa,” *Ilk. J. Ilm.*, vol. 12, no. 2, pp. 121–129, 2020, doi: 10.33096/ilkom.v12i2.542.121-129.
- [11] T. Suprawoto, “Klasifikasi data mahasiswa menggunakan metode k-means untuk menunjang pemilihan strategi pemasaran,” vol. 1, no. 1, pp. 12–18, 2016.
- [12] dkk Sugiono, “Pengelompokan Perilaku Mahasiswa Pada Perkuliahan E-Learning dengan K-Means Clustering,” *J. Kaji. Ilm.*, vol. 19, no. 2, pp. 126–133, 2019.
- [13] Z. I. Alfianti, M. A. Azis, A. Fauzi, F. T. Informasi, and U. B. Saranainformatika, “JurnalMantik,” vol. 4, no. 4, pp. 2336–2341, 2021.
- [14] S. Ramadani, I. Ambarita, and A. M. H. Pardede, “Metode K-Means Untuk Pengelompokan Masyarakat Miskin Dengan Menggunakan Jarak Kedekatan Manhattan City Dan Euclidean ( Studi Kasus Kota Binjai ),” *Inf. Syst. Dev.*, vol. 04, no. 2, pp. 15–29, 2019.
- [15] Satgas Covid-19 Karawang. “Pantauan kasus aktif Covid-19 untuk sebaran kecamatan” Internet: <http://covid19.karawangkab.go.id/data/>, tanggal di-update [20 Mei 2021].