



ANALISIS PERSEBARAN KASUS COVID-19 DI JAWA BARAT MENGUNAKAN METODE K-MEANS CLUSTERING

Elsa Ramadanti¹⁾, Muhamad Muslih²⁾

^{1,2}Program Studi Sistem Informasi, Universitas Nusa Putra

Jl. Raya Cibatu Cisaat No.21, Cibolang Kaler, Kec. Cisaat, Kab. Sukabumi, Jawa Barat 43155

Email : elsa.ramadanti_si18@nusaputra.ac.id¹⁾, muhamad.muslih@nusaputra.ac.id²⁾

*Korespondensi: muhamad.muslih@nusaputra.ac.id

ABSTRAK

Pada saat ini kasus COVID-19 di Indonesia belum juga mereda. Penambahan kasus COVID-19 hariannya belum mengalami penurunan signifikan. Persebaran kasus COVID-19 ini sudah menyebar hampir ke seluruh provinsi di Indonesia, salah satunya provinsi Jawa Barat. Terdapat 27 kabupaten/kota di Jawa Barat yang menjadi persebaran kasus COVID-19. Untuk memudahkan pemerintah daerah Jawa Barat dalam mengambil tindakan dalam upaya pencegahan penambahan persebaran kasus COVID-19 maka perlunya peneliti untuk menentukan tingkat persebaran kasus COVID-19 yang dibagi menjadi 3 cluster diantaranya yaitu cluster tinggi, sedang, dan rendah. Dalam penelitian ini peneliti menganalisis tingkat persebaran kasus COVID-19 menggunakan metode data mining dengan algoritma K-means Clustering. Untuk pengolahan data dengan K-means Clustering peneliti menggunakan aplikasi RapidMiner Studio 9.9. Hasil penelitian ini didapatkan bahwa untuk cluster tinggi terdapat 2 kabupaten/kota, cluster sedang 6 kabupaten/kota dan cluster rendah 19 kabupaten/kota tingkat persebaran kasus COVID-19 di Jawa Barat.

Kata Kunci: Analisis, COVID-19, Data Mining, K-Means Clustering.

ABSTRACT

At this time the cases of COVID-19 in Indonesia have not subsided. The daily addition of COVID-19 cases has not decreased significantly. The spread of COVID-19 cases has spread to almost all provinces in Indonesia, one of which is West Java province. There are 27 districts/cities in West Java that have spread COVID-19 cases. To facilitate the West Java regional government in taking action in an effort to prevent the addition of the spread of COVID-19 cases, it is necessary for researchers to determine the level of distribution of COVID-19 cases which is divided into 3 clusters, namely high, medium, and low clusters. In this study, researchers analyzed the spread of COVID-19 cases using a data mining method with the K-means Clustering algorithm. For data processing with K-means Clustering, researchers used the RapidMiner Studio 9.9 application. The results of this study showed that for the high cluster there were 2 regencies/cities, the medium cluster was 6 regencies/cities and the low cluster was 19 regencies/cities with the level of distribution of COVID-19 cases in West Java.

Keywords: : Analysis, COVID -19, Data Mining, K-Means Clustering. .

I. PENDAHULUAN

COVID-19 merupakan penyakit menular yang disebabkan oleh sindrom pernapasan akut virus SARS CoV-2. Gejala COVID-19 bervariasi, tetapi sering kali meliputi demam, batuk, sakit kepala, kelelahan, kesulitan bernapas, serta kehilangan penciuman dan perasa. Kasus pertama yang diketahui diidentifikasi di Wuhan, Cina pada Desember 2019. Penyakit ini telah menyebar ke seluruh dunia, menyebabkan pandemi yang berkelanjutan termasuk di Indonesia [1]. Penyebaran COVID-19 dapat meliputi berbagai media diantaranya melalui droplet saat seseorang batuk, bersin, berbicara, hingga bernapas. Selanjutnya penyebaran COVID-19 juga bisa terjadi melalui udara, permukaan yang terkontaminasi virus tersebut, dan limbah manusia sendiri. [2]



Kasus positif COVID-19 di Indonesia pertama kali dideteksi pada tanggal 2 Maret 2020, ketika dua orang terkonfirmasi tertular dari seorang warga negara Jepang. Setelah terkonfirmasi adanya kasus positif COVID-19 atau virus corona, pemerintah Indonesia melakukan berbagai cara atau upaya untuk menghindari terjadinya penularan COVID-19 yang berkepanjangan di Indonesia. Pada tanggal 9 April, pandemi sudah menyebar ke 34 provinsi, salah satunya provinsi Jawa Barat [3].

Virus Corona ini telah menyebar ke 27 kabupaten/kota di Jawa Barat. Adapun kabupaten/kota yang menjadi persebaran kasus COVID-19 antara lain yaitu kabupaten Bogor, Sukabumi, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Sukabumi, Bandung, Cirebon, Bekasi, Depok, Cimahi, Tasikmalaya dan Banjar. Hal ini tentu menyulitkan pemerintah untuk mengelompokkan tingkat persebaran kasus COVID-19 di Jawa Barat.

Pada penelitian ini, peneliti mengelompokkan persebaran kasus COVID-19 di Jawa Barat dengan mengelompokkan tingkat persebaran kasus COVID-19 ke dalam tiga cluster berdasarkan kabupaten/kota diantaranya yaitu cluster tinggi, sedang, dan rendah. Dimana hasil penelitian ini bisa digunakan oleh pemerintah Jawa Barat untuk mengambil kebijakan berdasarkan cluster persebaran kasus COVID-19 berdasarkan kabupaten/kota yang ada. Dalam pengelompokan cluster persebaran kasus COVID-19 di Jawa Barat ini, peneliti menggunakan metode Data Mining dengan algoritma K-Means Clustering. Data mining merupakan suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut [4]. Sedangkan K-means Clustering merupakan salah satu metode cluster analisis non hirarki yang berusaha untuk memisahkan objek atau data yang ada ke dalam satu atau lebih cluster berdasarkan karakteristik masing-masing cluster [5].

II. TINJAUAN PUSTAKA

Dalam suatu penelitian diperlukan dukungan hasil-hasil penelitian yang telah ada sebelumnya yang berkaitan dengan penelitian tersebut.

Dari penelitian Desy Noor Permata Saria, YL. Sukestiyarno (2021), Universitas Negeri Semarang tentang “Analisis Cluster dengan Metode K-Means pada Persebaran Kasus COVID-19 Berdasarkan Provinsi di Indonesia”. Dalam penelitiannya menggunakan metode K-Means Clustering untuk mengetahui tingkat persebaran kasus COVID-19 kategori tinggi, sedang, dan rendah pada masing-masing provinsi di Indonesia. Ada beberapa aspek yang diukur seperti jumlah penduduk, kepadatan penduduk, kasus positif terinfeksi COVID-19, pasien yang sembuh, dan pasien yang meninggal dunia dan data tersebut dianalisis menggunakan software SPSS [6].

Penelitian lain dari Dina Tri Utari (2021), Universitas Islam Indonesia, tentang “Analisis Karakteristik Wilayah Transmisi Covid-19 dengan Menggunakan Metode K-Means Clustering”. Penelitian ini bertujuan untuk melihat kejadian kasus positif COVID-19 berdasarkan jenis transmisi COVID-19 yang terjadi di seluruh wilayah Bali. Hasil penelitian ini menunjukkan bahwa berdasarkan sumber transmisi dapat dikelompokkan menjadi empat cluster yang memiliki karakteristik masing-masing. Untuk cluster 1,2 dan 3 antara lain pembatasan transmisi lokal dan perjalanan dalam negeri. Sedangkan pembatasan transmisi lokal serta perjalanan luar negeri pada cluster 4 [7].

Penelitian selanjutnya dari Darmansah, Ni Wayan Wardani (2021), Institut Teknologi Telkom Purwokerto, tentang “Analisis Pesebaran Penularan Virus Corona Di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering”. Dalam penelitiannya menunjukkan tentang menentukan tingkat pesebaran penularan virus corona di Jawa Tengah dengan metode K-means Clustering yang dibagi menjadi 3 cluster. Cluster pertama yaitu C0 dengan kategori rendah, C1 dengan kategori sedang dan C2 dengan kategori pesebaran tinggi. Kemudian dalam pengolahan data peneliti menggunakan aplikasi Rapidminer Studio 7.6 [8].

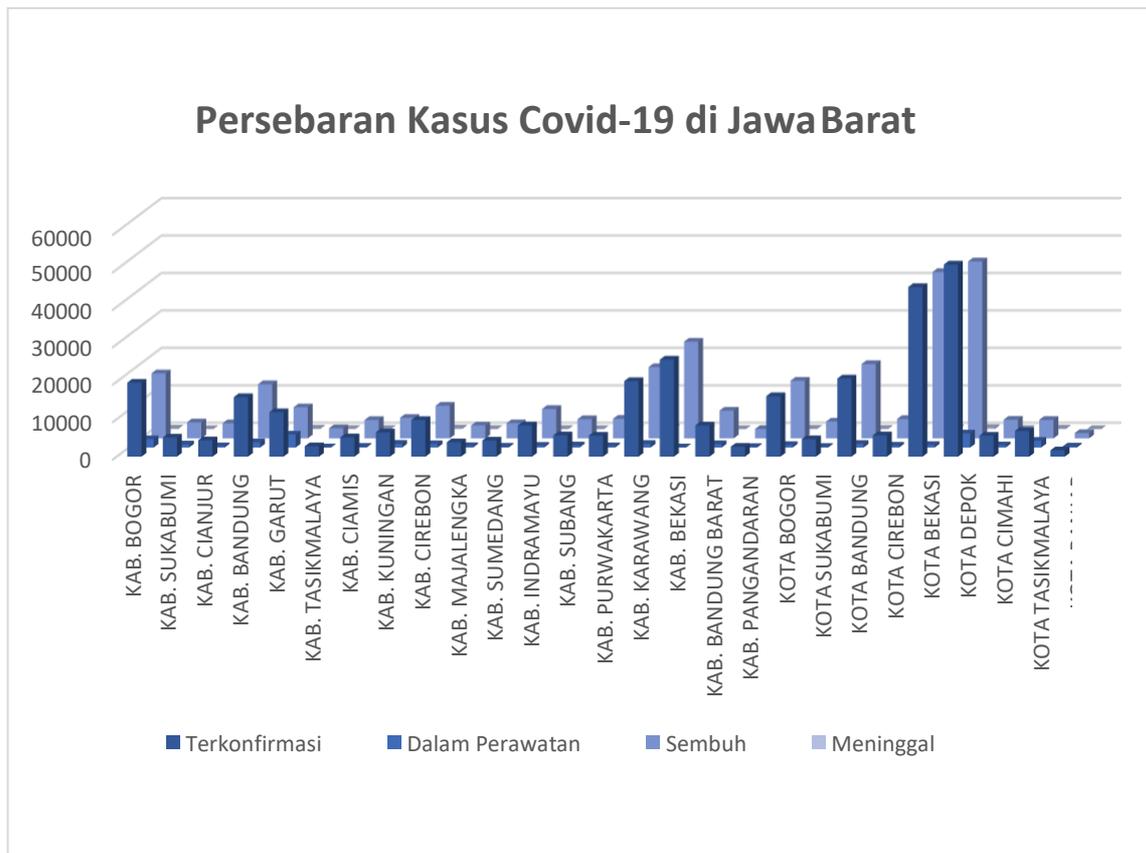
Dari ketiga sumber penelitian tersebut dengan penelitian ini menunjukkan bahwa dalam penelitiannya menggunakan metode yang sama yaitu K-means Clustering untuk mengelompokkan suatu data ke dalam beberapa tingkat cluster yang telah ditentukan.

III. METODOLOGI PENELITIAN

Adapun tujuan dari penelitian ini adalah untuk mengelompokkan data persebaran kasus covid-19 di Jawa Barat menggunakan metode K-Means Clustering. Untuk melakukan clustering menggunakan metode k-means ada beberapa langkah yang harus dilakukan adalah sebagai berikut :

2.1 Analisis Data

Data yang digunakan dalam penelitian ini adalah jumlah kasus positif Covid-19 pada setiap wilayah kabupaten/kota di Jawa Barat dari website <https://pikobar.jabarprov.go.id> dimana data ini merupakan data sekunder yaitu data persebaran kasus covid-19 per 11 Juni 2021 yang terdiri dari 27 kabupaten/kota di Jawa Barat [9].



Gambar 1. Persebaran Kasus Covid-19 di Jawa Barat

2.2 Pengolahan Data

Adapun langkah-langkah proses algoritma K-Means Clustering adalah sebagai berikut:

- Menentukan jumlah data yang akan di cluster. Adapun cluster yang akan dibuat adalah 3 cluster yaitu tinggi, sedang dan rendah.
- Menentukan titik pusat cluster centroid awal (iterasi 1) yang telah ditentukan secara random/acak. Titik pusat cluster yang didapatkan dapat dilihat pada tabel berikut:

Tabel 1. Titik Pusat Awal Tiap Cluster

Cluster	KOTA/KAB	Terkonfirmasi	Dalam Perawatan	Sembuh	Meninggal
C1	KOTA DEPOK	51316	3748	47246	322
C2	KAB. GARUT	11891	3461	8287	143
C3	KOTA BANJAR	1694	261	1409	24

- c. Menghitung jarak setiap data ke pusat cluster antara objek ke centroid terdekat. Centroid terdekat akan menjadi cluster yang diikuti oleh data tersebut. Adapun rumus persamaannya sebagai berikut:

$$d(P, Q) = \sqrt{\sum_{j=1}^p (x_j(P) - x_j(Q))^2}$$

Keterangan :

D = titik

dokumen P =

data record

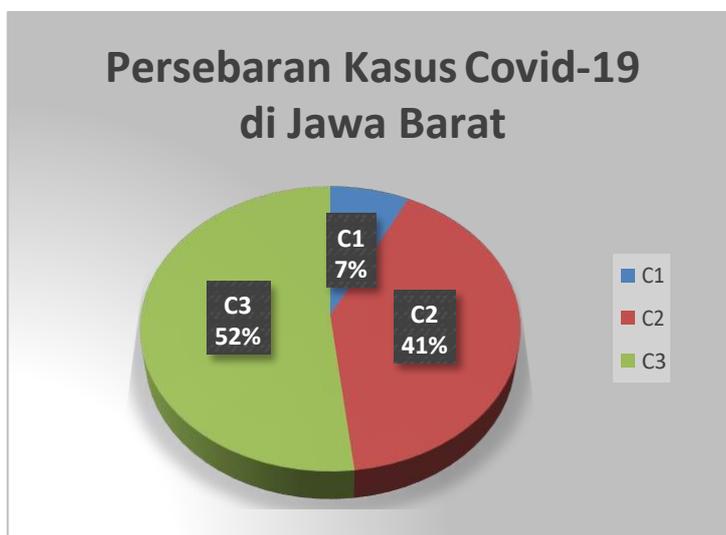
Q = data centroid

Tabel 2. Hasil Perhitungan Jarak Terdekat Cluster Iterasi 1

No.	C1	C2	C3	Jarak Terdekat	Cluster
1	43479,58	12081,84	24190,21	12081,84398	Cluster 2
2	63160,28	8297,661	4502,932	4502,931712	Cluster 3
3	63897,69	9183,665	3748,99	3748,990264	Cluster 3
4	48324,69	7639,24	19319,66	7639,240407	Cluster 2
5	55427,87	0	12709,83	0	Cluster 2
..
27	67643,34	12709,83	0	0	Cluster 3

Tabel 3. Hasil Cluster Iterasi 1

Cluster	Provinsi	Hasil
C1	23,24	2
C2	1,4,5,9,12,15,16,17,19,21,26	11
C3	2,3,6,7,8,10,11,13,14,18,22,25	14



Gambar 2. Clustering Data Iterasi 1

- d. Setelah semua data ditempatkan kedalam cluster yang terdekat, kemudian dihitung kembali ke titik pusat cluster yang baru berdasarkan rata-rata anggota yang ada pada cluster tersebut. dengan rumus berikut:

$$C = \frac{\sum m}{n}$$

Dimana:

C: centroid data

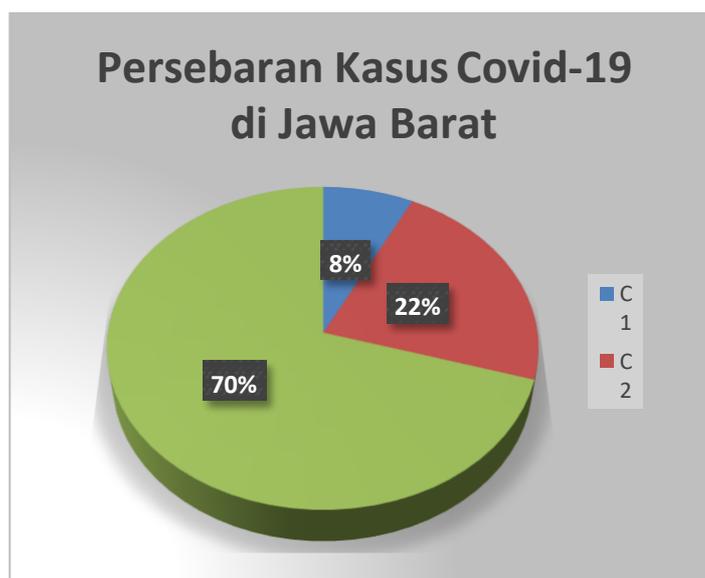
m : anggota data yang termasuk kedalam centroid

tertentu n : jumlah data yang menjadi anggota centroid tertentu

- e. Proses K-Means akan terus beriterasi sampai pengelompokan data sama dengan pengelompokan data iterasi sebelumnya. Dengan kata lain, proses akan terus melakukan iterasi sampai data pada iterasi terakhir sama dengan iterasi sebelumnya. Sampai proses iterasi dalam penelitian ini berhenti pada iterasi ke 4 yang dapat dilihat hasil akhirnya pada tabel berikut.

Tabel 4. Hasil Cluster Iterasi 4

Cluster	Provinsi	Hasil
C1	23,24	2
C2	1,4,15,16,19,21	6
C3	2,3,5,6,7,8,9,10,11,12,13,14,17,18,20,22,25,26,27	19



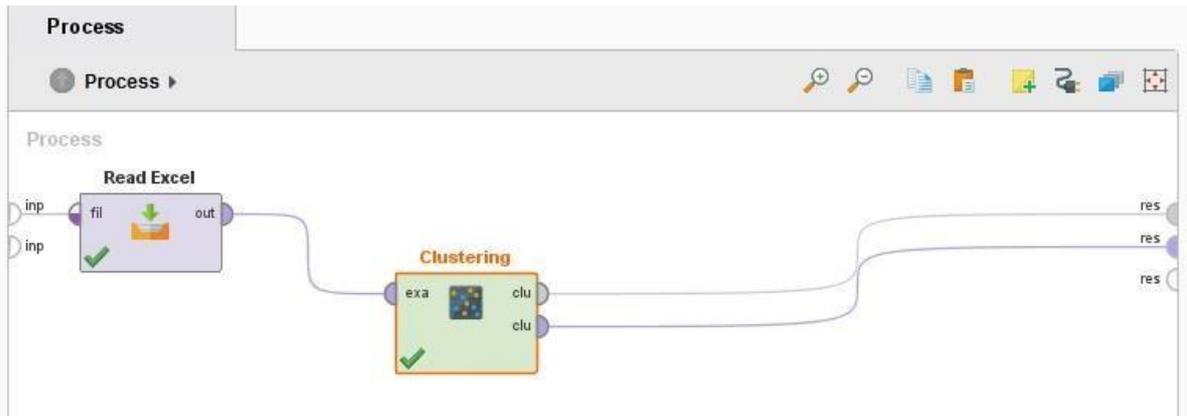
Gambar 3. Clustering Data Iterasi 4

Pada iterasi ke 4, titik pusat dari setiap cluster sudah tidak berubah dan tidak ada lagi data yang berpindah dari satu cluster ke cluster yang lain.

IV. HASIL DAN PEMBAHASAN

3.1 Proses K-Means Clustering pada RapidMiner

Adapun pengolahan data persebaran kasus covid-19 di Jawa Barat dengan menggunakan metode k-means clustering menggunakan software RapidMiner dapat dilihat pada gambar berikut:



Gambar 4. Proses Clustering pada RapidMiner

3.2 Hasil Clustering

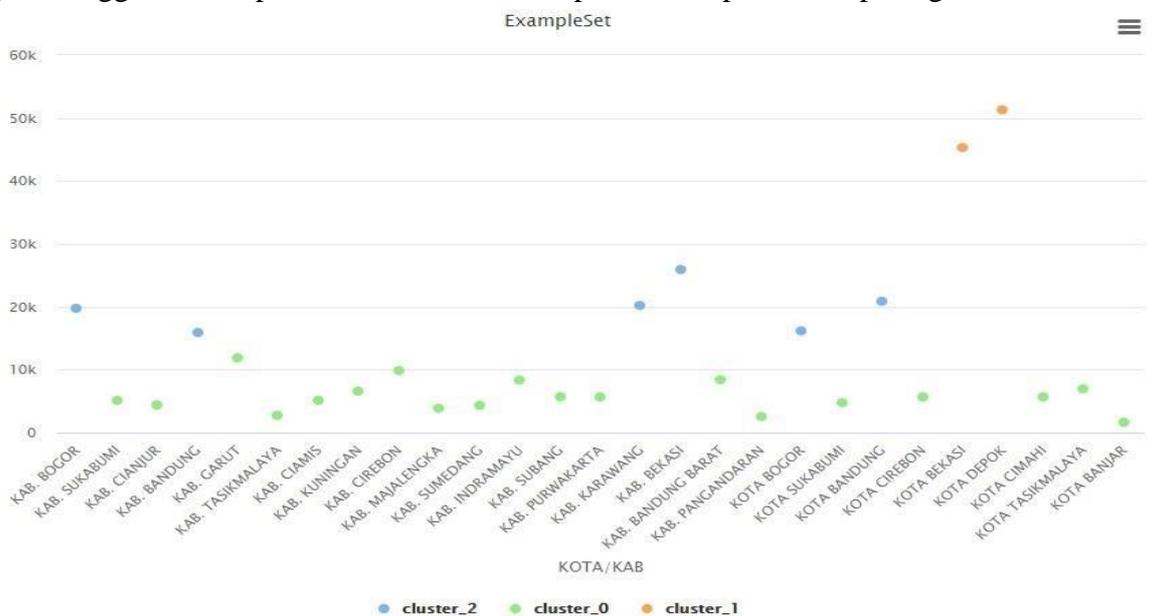
Setelah dilakukan perhitungan dan proses K-means Clustering, maka di dapatkan hasil cluster model seperti berikut ini:

Cluster Model

```
Cluster 0: 19 items
Cluster 1: 2 items
Cluster 2: 6 items
Total number of items: 27
```

Gambar 5. Hasil Clustering

Hasil Penyebaran cluster 0, cluster 1, dan cluster 2 sebanyak 27 data pada proses k-means clustering dengan menggunakan rapidminer, untuk 3 kelompok data dapat dilihat pada gambar berikut :



Gambar 6. Hasil Akhir Pengelompokan Rapidminer



Pada Gambar 6 dapat dilihat dimana cluster 1 pada RapidMiner adalah cluster tinggi berjumlah 2 kabupaten/kota yang dapat dilihat pada titik merah, cluster sedang yaitu cluster 2 berjumlah 6 kabupaten/kota pada titik hijau, sedangkan cluster rendah yaitu cluster 0 berjumlah 19 kabupaten/kota pada titik biru. Berdasarkan dari penjelasan di atas mengenai tahap-tahap pengolahan data serta hasil yang telah ditampilkan maka akan dibahas mengenai keterkaitan dari hasil yang didapat dari RapidMiner dan Microsoft excel menjelaskan bahwa hasil dari perhitungan manual algoritma k-means dan Microsoft excel data memiliki nilai yang sama dari jumlah setiap cluster yaitu untuk cluster tinggi, cluster sedang dan cluster rendah.

V. KESIMPULAN

Berdasarkan hasil analisis dan pengujian yang dilakukan menggunakan metode data mining dengan algoritma K-Means Clustering terhadap data persebaran kasus Covid-19 di Jawa Barat yang diperoleh dari melalui situs <https://pikobar.jabarprov.go.id>. Maka dapat disimpulkan bahwa hasil penelitian ini adalah sebagai berikut. Analisa tingkat persebaran kasus Covid-19 di Jawa Barat menggunakan algoritma K-Means Clustering menggunakan aplikasi Microsoft Excel dan Rapidminer dalam perhitungan dan pengolahan data nya untuk menentukan jumlah cluster. Cluster dibagi menjadi 3 cluster yaitu cluster tinggi, sedang dan rendah.

Data yang digunakan berjumlah sebanyak 27 data hal ini berdasarkan jumlah kabupaten/ kota se-Jawa Barat. Yang termasuk ke dalam cluster tinggi penyebaran kasus covid-19 terdiri dari 2 kabupaten/kota yaitu kota Bekasi dan Depok. Selanjutnya yang tergolong ke dalam cluster sedang terdapat 6 kabupaten/kota diantaranya kabupaten Bogor, Bandung, Karawang, Bekasi, kota Bogor dan Bandung. Terdapat 19 kabupaten/kota yang termasuk ke dalam cluster rendah yaitu kabupaten Sukabumi, Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Bandung Barat, Pangandaran, kota Sukabumi, Cirebon, Cimahi, Tasikmalaya dan Banjar.

Hasil dari proses data mining ini dapat digunakan sebagai pertimbangan dalam mengambil keputusan lebih lanjut tentang kebijakan yang di lakukan dinas kesehatan provinsi Jawa Barat dalam menangani tingkat persebaran kasus covid-19 di kabupaten/kota yang ada di Jawa Barat, sehingga kualitas penanganan pencegahan covid-19 semakin lebih baik.

DAFTAR PUSTAKA

- [1] Wikipedia, "COVID-19," Wikipedia, [Online]. Available: <https://en.wikipedia.org/wiki/COVID-19>. [Accessed 21 Juni 2021].
- [2] S. O. Alam, "Berbagai Cara Penyebaran Virus Corona COVID-19 Menurut WHO, Apa Saja?," *detikHealth*, 6 Agustus 2020. [Online]. Available: <https://health.detik.com/berita-detikhealth>. [Accessed 20 Juni 2021].
- [3] Wikipedia, "Pandemi COVID-19 di Indonesia," Wikipedia, [Online]. Available: https://id.wikipedia.org/wiki/Pandemi_COVID-19_di_Indonesia. [Accessed 20 Juni 2021].
- [4] N. Rofiqo, A. P. Windarto and D. Hartama, "Penerapan Clustering pada Penduduk yang Mempunyai Keluhan Kesehatan Dengan Datamining K-Means," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. II, no. 1, p. 216–223, 2018.
- [5] M. G. Sadewo, A. Eriza, A. P. Windarto and D. Hartama, "Algoritma K-Means Dalam Mengelompokkan Desa/Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi," *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, p. 754–761, 2019.
- [6] D. N. P. Saria and Y. Sukestiyarno, "Analisis Cluster dengan Metode K-Means pada Persebaran Kasus COVID-19 Berdasarkan Provinsi di Indonesia," *PRISMA, Prosiding Seminar Nasional Matematika*, pp. 602-610, 2021.



- [7] D. T. Utari, "Analisis Karakteristik Wilayah Transmisi Covid-19 dengan Menggunakan Metode K-Means Clustering," *Jurnal Media Teknik & Sistem Industri* , vol. V, no. 1, p. 25 – 32 , 2021.
- [8] D. and N. W. Wardani, "Analisis Pesebaran Penularan Virus Corona Di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering," *Jurnal Teknik Informatika dan Sistem Informasi* , vol. VIII, no. 1, pp. 105-117 , 2021.
- [9] "Sebaran Kasus Covid-19 di Jawa Barat," Pusat Informasi dan Koordinasi COVID-19 Provinsi Jawa Barat, 11 Juni 2021. [Online]. Available: <https://pikobar.jabarprov.go.id/>. [Accessed 11 Juni 2021].