

Analisis Penerapan Data Mining Terhadap Kasus Positif Covid-19 Menggunakan Metode K-Means Clustering

Ridhan Azhari*, Dedy Hartama, Muhammad Ridwan Lubis, Della Patricia Nasution, Agus Perdana Windarto

STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: ^{1,*}ridhanazhari11@gmail.com, ²dedyhartama@amiktunasbangsa.ac.id, ³Ridwan@amiktunasbangsa.ac.id,

⁴dellafatrician@gmail.com, ⁵agus.perdana@amiktunasbangsa.ac.id

Email Penulis Korespondensi: ridhanazhari11@gmail.com

Abstrak—Penelitian ini memiliki masalah seperti belum adanya penggunaan algoritma K-means clustering terhadap data kasus positif COVID-19 di seluruh provinsi Indonesia. Tujuan dari penelitian skripsi ini yakni menerapkan metode K-means clustering dalam mencari jarak terdekat untuk menghasilkan cluster terendah dan tertinggi pada data kasus positif COVID-19 di seluruh provinsi Indonesia. Kmeans merupakan salah satu algoritma dalam teknik Clustering non-hierarki yang berusaha membagi sebuah data yang ada ke dalam bentuk satu atau lebih cluster. Hasil yang diperoleh dari metode k-means clustering yaitu menghasilkan 2 cluster yaitu cluster terendah C1 = 30 item Sedangkan untuk cluster tertinggi C2 = 4 item. Penelitian ini dapat digunakan sebagai referensi dan dapat dikembangkan lagi dengan metode atau algoritma clustering lainnya seperti k-medoid agar mendapat perbandingan hasil serta langkah-langkah penggunaan algoritma yang berhubungan dengan clustering.

Kata Kunci: COVID-19; Clustering; Data Mining; K-Means

Abstract—This study has problems such as the absence of the use of the K-means clustering algorithm for data on positive COVID-19 cases in the Indonesian province. The purpose of this study is to apply the K-means clustering method in finding the closest distance to produce the lowest and highest clusters of data on positive COVID-19 cases in the Indonesian province. K-means is one of the algorithms in the non-hierarchical Clustering technique that tries to partition the existing data in the form of one or more clusters. The results obtained from the k-means clustering method produced 2 clusters, namely the lowest cluster C1 = 30 items while the highest cluster C2 = 4 items. This research can be used as a reference and can be further developed with other clustering methods or algorithms such as k-medoid in order to get a comparison of results and steps to use algorithms related to clustering.

Keywords: COVID-19; Clustering; Data Mining; K-Means

1. PENDAHULUAN

Pada awal 2020, dunia dikejutkan dengan mewabahnya pneumonia baru yang bermula dari Wuhan, Provinsi Hubei yang kemudian menyebar dengan cepat ke lebih dari 190 negara dan teritori. Wabah ini diberi nama coronavirus disease 2019 (COVID-19) yang disebabkan oleh Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2). Penyebaran penyakit ini telah memberikan dampak luas secara sosial dan ekonomi. Masih banyak kontroversi seputar penyakit ini, termasuk dalam aspek penegakkan diagnosis, tata laksana, hingga pencegahan. Oleh karena itu, kami melakukan telaah terhadap studi-studi terkait COVID-19 yang telah banyak dipublikasikan sejak awal 2020 lalu sampai dengan akhir Maret 2020[1], [2].

Virus ini dapat menular dengan cepat bahkan sangat cepat yang mana penyebarannya langsung berdampak pada seluruh dunia termasuk Indonesia hanya dalam waktu beberapa bulan saja. Hal tersebut membuat beberapa negara melakukan kebijakan lockdown untuk mencegah virus corona semakin meluas termasuk Indonesia. Banyaknya jumlah masyarakat Indonesia membuat virus ini menyebar sangat cepat, tepatnya pada akhir tahun 2020 Indonesia sudah terkonfirmasi 743.198 kasus positif COVID-19, dan pada tahun 2021 kasus COVID-19 terus naik melambung tinggi sehingga menjadi negara dengan jumlah kasus positif COVID-19 terbanyak di Asia Tenggara yaitu mencapai 4.211.4602 kasus positif COVID-19 pada akhir tahun tepatnya pada tanggal 28 September 2021 data tersebut diperoleh dari <http://covid19.go.id>[3], [4].

Pada data kasus positif COVID-19 yang sudah terkonfirmasi di atas memiliki masalah seperti belum adanya penggunaan algoritma K-means clustering terhadap data kasus positif COVID-19 di provinsi Indonesia dalam menemukan jarak terdekat untuk menghasilkan sebuah cluster (kelompok) menggunakan algoritma K-means clustering[5], [6].

K-means merupakan salah satu algoritma dalam teknik Clustering nonhierarki yang berusaha memartisi data yang ada ke dalam bentuk satu atau lebih cluster. Algoritma ini memartisi data ke dalam cluster sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik berbeda dikelompokkan ke dalam cluster yang lain. Melalui data yang diterbitkan dari website Satgas COVID-19 kita dapat mengelola data tersebut menggunakan algoritma K-means Clustering untuk menghasilkan cluster terendah dan tertinggi melalui data kasus positif COVID-19 di provinsi Indonesia[7]–[9].

“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[10].

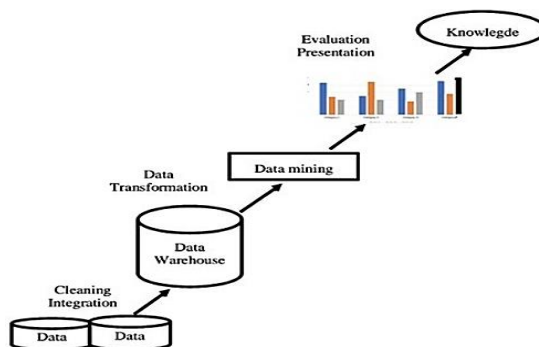
Berdasarkan hasil analisis dan pengujian yang dilakukan menggunakan metode data mining dengan algoritma K-Means Clustering terhadap data persebaran kasus COVID-19 di Jawa Barat yang diperoleh dari melalui situs <https://pikobar.jabarprov.go.id>. Maka dapat disimpulkan bahwa hasil penelitian ini adalah 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.

2. METODOLOGI PENELITIAN

2.1 Data Mining

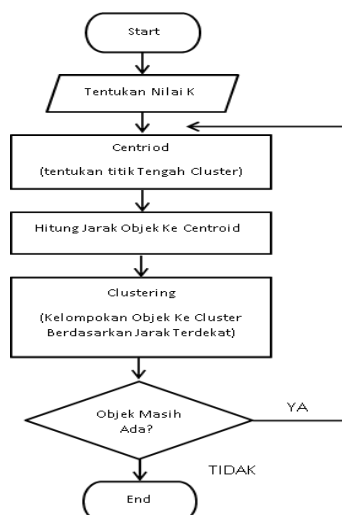
Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran komputer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. Data mining merupakan langkah analisis terhadap sekumpulan data yang umumnya berukuran besar untuk mendapatkan hubungan antar data tersebut dan merangkumnya dalam bentuk yang mudah dipahami dan digunakan. Data mining merupakan proses *iteratif* dan *interaktif* untuk menemukan pola atau model baru yang sah (sempurna), bermanfaat dan dapat dimengerti dalam suatu *database* yang sangat besar (*massive databases*). Data mining sering disebut juga dengan KDD (Knowledge Discovery Database) yaitu keseluruhan proses ekstraksi atau identifikasi pola, pengetahuan dan informasi potensi dari sekumpulan data besar. Pengetahuan dan informasi yang dihasilkan dari KDD bersifat sah, baru, mudah dimengerti serta bermanfaat[11]–[13]. Berikut adalah gambar tahapan proses KDD, seperti pada gambar 1.



Gambar 1. Tahapan Proses KDD

2.2 Clustering

Clustering adalah proses pengelompokan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan antar objek dengan prinsip untuk memaksimalkan kesamaan antar anggota satu kelas dan menimbulkan kesamaan antar *cluster*. Didalam menentukan sebuah *cluster* berdasarkan data yang tersedia, “dibutuhkan sebuah *Flowchart* untuk mengetahui ketentuan alur perhitungan sebagai alur untuk menemukan suatu hasil dalam penerapan sebuah *cluster*”, seperti pada gambar 1[14], [15].



Gambar 2. Flowchart K-Means

Dari gambar 2 menjelaskan bagaimana tahapan dalam menentukan Nilai *cluster* yang ada secara acak. Menghitung titik *centroid* tiap objek pada suatu data yang digunakan. Mencari data dengan menggunakan rata-rata jarak terdekat jika tidak ada perubahan ulangi sampai nilai *centroid* berubah menjadi *threshold*.

2.3 Algoritma K-Means

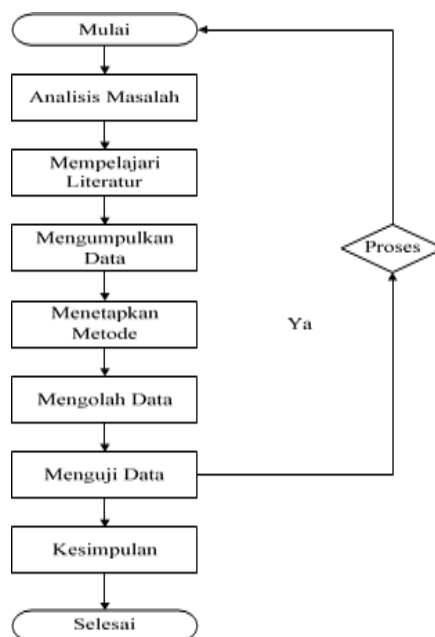
Algoritma *K-means clustering* digunakan di dalam data mining untuk mengelompokkan suatu data menjadi bagian partisi-partisi yang diinginkan. Ada beberapa alur yang harus dilakukan agar data tersebut dapat dianalisis oleh algoritma *K-means clustering*. *K-means* merupakan salah satu algoritma dalam teknik *Clustering* non- hierarki yang berusaha memartisi data yang ada ke dalam bentuk satu atau lebih *cluster*. Algoritma ini memartisi data kedalam *cluster* sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik berbeda dikelompokkan kedalam *cluster* yang lain[16]–[18]

2.4 RapidMiner Studio

Rapidminer merupakan perangkat lunak yang bersifat terbuka. *Rapidminer* ini sebuah alat yang menganalisis kerangka kerja tanpa perlu menulis kode apa pun dalam memberikan wawasan terhadap pengguna, (siregar, 2017). *Rapidminer* ditulis dalam bahasa pemrograman java dan juga mengintegrasikan proyek data ke dalam data mining weka dan statistika *rapidminer* mendapatkan peringkat satu sebagai *tools* data mining untuk proyek nyata pada poll oleh Kdnuggets pada senuah koran kabar tahun 2010-2011. Sebelumnya, *rapidminer* bernama YALE (*Yet Another Learning Environmet*). Dimana versi awalnya di kembangkan pada tahun 2001 oleh Ralfklingsenber, Igo Mieswa, dan Simon Fischer di Artificial Intelligence Unit dari University Of Dortmund[19], [20].

2.5 Rancangan Penelitian

Rancangan penelitian digunakan untuk menjelaskan proses kerja dalam menyelesaikan masalah penelitian. Alur penelitian ini akan digambarkan Berdasarkan *flowchart* dibawah ini seperti pada gambar 3.



Gambar 3. Rancangan Penelitian

Pada gambar 3 di atas menjelaskan untuk menentukan kelompok dalam kasus positif COVID-19 disetiap provinsi Indonesia. Dengan menggunakan metode *K-means Clustering* yang dihasilkan terdapat keterangan antara lain sebagai berikut.

1. Analisis Masalah, Menganalisis masalah yang terkait dengan mencari jarak terdekat untuk menghasilkan *cluster* terendah dan tertinggi pada data kasus positif COVID-19 diprovinsi indonesia. Kriteria yang digunakann kasus terendah dan tertinggi pada tahun 2020 dan tahun 2021.
2. Mempelajari Literatur, Penelitian ini harus didasari rujukan yang digunakan untuk mendapatkan sebuah informasi dalam penelitian.
3. Menetapkan Metode, Menetapkan metode untuk mencari solusi masalah. Pada penelitian ini digunakan teknik data mining dengan metode *k-means clustering*.
4. Mengumpulkan Data, Data yang dikumpulkan didapat dari website Satgas COVID-19 Indonesia, pasa tanggal 27 Desember 2021.
5. Mengolah Data, Melakukan pengolahan dengan Data Transformation Service (DTS) yaitu dari bentuk data mentah dipindahkan kedalam bentuk data olahan siap pakai dengan Ms.Excel 2010 (.xlsx).

6. Menguji Data, Pengujian data dilakukan menggunakan Software RapidMiner Studio versi 5.3 .
7. Kesimpulan, Kesimpulan yang didapat dalam menentukan nilai minimum dan maximum untuk mencari jarak terdekat agar menghasilkan cluster terendah dan tertinggi dalam sebuah penelitian menggunakan algoritma *K-means clustering*.

Hasil pengamatan kemudian dibuat percobaan yang mendukung, selanjutnya dilakukan eksperimen data dengan menggunakan *Rapidminer Studio versi 5.3* yang merupakan *software open source*, untuk melakukan pengelompokan terhadap data kasus positif COVID-19 di Indonesia data tersebut berbentuk .xls. Terlihat pada Tabel 1 dan Tabel 2

Tabel 1. Data Awal Kasus Positif COVID-19 Tahun 2020

Provinsi	maret	april	mei	juni	juli	agustus	september	oktober	november	desember
Aceh	5	5	10	60	330	1223	2921	2851	931	454
Bali	12	205	244	1023	1919	1800	3670	2891	2121	3714
Banten	79	277	463	601	375	1069	2756	3849	3552	5113
Bengkulu	1	11	79	34	92	126	350	388	719	1780
DKI Jakarta	371	3390	3255	4163	10183	18685	33653	31861	30817	46873
D.I Yogyakarta	11	83	141	77	361	752	1218	1192	2115	6192
Gorontalo	0	15	79	155	766	1025	721	264	89	731
Jambi	2	30	65	20	45	140	211	717	710	1284
Jawa Barat	159	634	1805	967	3285	5038	11561	15062	18439	33450
Jawa Tengah	118	844	917	3504	7143	4474	8469	11459	20883	24868
Jawa Timur	106	887	3861	6968	9898	11547	10256	8924	9450	22494
Kalimantan Barat	8	50	131	132	66	258	337	673	746	713
Kalimantan Selatan	7	163	747	2231	2950	2190	2060	1511	1356	2104
Kalimantan Tengah	7	138	261	488	835	819	1097	699	1595	3808
Kalimantan Timur	16	116	163	223	908	2817	4407	5440	5452	7459
Kalimantan Utara	2	98	65	41	77	101	193	245	639	2329
Kepulauan Bangka Belitung	2	9	36	106	41	46	122	222	421	1335
Kepulauan Riau	5	84	106	98	158	540	1243	1774	1449	1537
Lampung	7	39	88	56	65	140	499	931	2003	2453
Maluku	1	22	200	519	351	764	958	1037	788	1102
Maluku Utara	1	39	113	574	813	320	205	134	175	398
Nusa Tenggara Barat	2	228	406	598	831	677	574	656	696	985
Nusa Tenggara Timur	0	3	89	21	32	32	247	257	519	958
Papua	9	196	466	1042	1346	737	2525	2716	977	3195
Papua Barat	2	35	131	71	194	374	1327	2015	1139	703
Riau	3	39	77	109	218	1401	5783	7168	5254	4892
Sulawesi Barat	1	41	50	23	115	165	399	235	291	471
Sulawesi Selatan	51	440	1044	3519	4323	2469	3664	2790	2240	10395
Sulawesi Tengah	2	45	81	58	21	36	167	447	1003	1692
Sulawesi Tenggara	3	59	182	119	419	784	1290	2094	1559	1404
Sulawesi Utara	1	44	294	770	1471	1268	639	922	1466	2744
Sumatera Barat	8	140	419	159	222	1209	4124	8249	5578	3423
Sumatera Selatan	5	145	831	1068	1368	1037	1641	1754	1625	2363
Sumatera Utara	13	104	292	1140	2382	2896	3484	2787	2487	2564

Tabel 2. Data Awal Kasus Positif COVID-19 Tahun 2021

Provinsi	Januari	Februari	Maret	April	Mei	Juni	Juli	Agustus	September	Oktober	November	Desember
Aceh	482	313	351	1153	3972	4229	3800	9890	4796	583	97	2
Bali	8559	8063	5469	4987	2588	2959	2610	3051	5746	1289	342	20
Banten	8040	3278	8287	9584	2133	7646	5560	1628	2458	868	341	17
Bengkulu	976	354	582	1321	1292	2057	7686	4653	551	70	12	1
DKI Jakarta	85981	70008	42326	26565	21440	113408	271192	35929	7182	3776	2406	98
DI Yogyakarta	9670	5998	5649	6044	5230	15714	57373	32085	4911	1020	901	19
Gorontalo	467	486	345	255	112	341	2075	3292	522	91	7	0

Provinsi	Januar i	Febru ar i	Maret	Apri l	Mei	Juni	Juli	Agu st us	Septe m ber	Okto ber	Nov e mber	Dese mber
Jambi	1353	908	737	1445	1942	3249	7392	7902	1392	194	26	1
Jawa Barat	62841	60810	37878	30930	32730	69735	220752	85031	11847	3059	1964	141
Jawa Tengah	43667	27682	15461	15411	16215	54087	129148	85103	12689	3101	1275	50
Jawa Timur	28497	16545	10083	8196	7038	18254	135384	74823	12225	2793	1165	65
Kalima ntan Barat	814	708	1262	1836	3116	3815	11077	10251	3963	1167	401	8
Kalimantan Selatan	2773	3761	6045	5076	1801	1317	11921	17918	3509	401	91	11
Kalimantan Tengah	2236	1875	3389	2831	2492	3178	9005	8618	1799	376	78	5
Kalima ntan Timur	14138	14180	8269	4964	2968	5636	40272	33195	6113	1123	298	13
Kalimantan Utara	3246	2654	1438	613	525	780	8375	10668	2914	755	122	1
Kepulauan Bangka	2330	2679	2017	4082	4851	2894	12184	13727	4107	791	267	23
Belitung	1104	603	483	1990	5652	9097	18855	7425	1426	215	37	2
Kepulauan Riau	3733	2526	1484	1956	2222	3475	13094	11900	2390	512	114	9
Lampung	704	558	368	178	214	748	4907	910	215	46	16	1
Maluku	681	530	346	89	101	705	4815	1752	166	124	22	0
Maluku Utara	1910	1166	854	857	1514	1044	6701	6151	1570	292	68	2
Nusa Tenggara Barat	2738	4342	3168	1938	1907	2156	20690	20343	3267	907	390	19
Nusa Tenggara Timur	2021	1812	2547	776	96	520	6054	5241	1467	455	145	4
Papua	748	696	886	633	374	1423	7801	3675	684	190	280	11
Papua Barat	4042	2390	3354	9525	15793	10447	26818	26220	4174	727	345	12
Riau	1791	1491	142	82	54	331	2539	3185	603	163	42	1
Sulawesi Barat	17214	7937	3433	1865	681	2056	19481	20926	3945	1096	230	2
Sulawesi Selatan	4241	2232	1196	1961	614	749	9176	19700	3820	672	119	16
Sulawesi Tengah	1594	542	224	153	100	854	5001	3278	377	94	47	3
Sulawesi Tenggara	3783	1548	363	285	122	490	8285	7774	1858	438	85	2
Sulawesi Utara	3515	2203	2546	5203	7327	6928	19687	15841	2562	495	73	5
Sumatra Barat	2484	1604	1813	2721	3884	4431	18263	10943	1646	253	69	5
Sumatra Selatan	2716	3663	2830	2041	2594	4280	24248	35787	8402	1121	227	8
Sumatra Utara												

Tabel 1 dan Tabel 2 merupakan data mentah yang belum melakukan *transformasi* data. Dimana data yang sudah di transformasi nantinya akan dipakai untuk melakukan proses perhitungan menggunakan algoritma *K-means Clustering* terhadap data kasus positif COVID-19, menggunakan 2 cluster yaitu clustert 0 =1, cluster ke 1=2.

3. HASIL DAN PEMBAHASAN

3.1 Pengolahan Data Algoritma K-Means

Pengolahan Data Algoritma K-means pada tahapan ini dilakukanlah analisis terhadap data kasus positif COVID-19 berdasarkan data yang diperoleh dari satgas COVID-19 Indonesia. Data tersebut akan dikelompokkan dan di proses dengan perhitungan manual *Algoritma K-means Clustering* dari perhitungan yang menggunakan microsoft Excel untuk melihat keakuratan hasil yang didapat perhitungan menggunakan *software rapidminer studio versi 5.3*

1. Menentukan data yang akan di cluster

Sampel data yang akan digunakan dalam proses *clustering* pada data kasus positif COVID-19 berdasarkan jumlah kasus positif COVID-19 diprovinsi Indonesia.

Tabel 3. Data Kasus Positif COVID-19 di Provinsi Indonesia

Provinsi	Tahun 2020	Tahun 2021	Jumlah
Aceh	8.573	29.737	38.310
Sumatra Utara	17.623	88.203	105.826
Sumatra Barat	22.458	67.310	89.768
Riau	23.727	104.718	128.445
Jambi	2.936	26.805	29.741
Sumatra Selatan	10.847	49.019	59.866
Bengkulu	3.076	20.015	23.091
Lampung	5.332	44.231	49.563
Kepulauan Bangka Belitung	1.697	50.297	51.994
Kepulauan Riau	6.528	47.316	53.844
Banten	16.163	116.141	132.304
DKI Jakarta	163.111	638.169	801.280
Jawa Barat	73.946	631.781	705.727
Jawa Tengah	71.531	413.525	485.056
DI Yogyakarta	9.497	146.331	155.828
Jawa Timur	75.274	322.930	398.204
Bali	16.176	97.668	113.844
Nusa Tenggara Barat	5.238	22.743	27.981
Nusa Tenggara Timur	1.808	61.807	63.615
Kalimantan Barat	2.950	38.150	41.100
Kalimantan Tengah	8.617	36.919	45.536
Kalimantan Selatan	14.586	55.236	69.822
Kalimantan Timur	24.081	133.843	157.924
Kalimantan Utara	2.806	32.946	35.752
Sulawesi Utara	8.937	25.678	34.615
Sulawesi Tengah	2.773	44.229	47.002
Sulawesi Selatan	25.763	83.907	109.670
Sulawesi Tenggara	7.550	12.573	20.123
Gorontalo	3.368	8.462	11.830
Sulawesi Barat	1.707	10.615	12.322
Maluku	5.448	9.122	14.570
Maluku Utara	2.654	9.424	12.078
Papua Barat	5.748	17.341	23.089
Papua	12.759	21.433	34.192

2. Menentukan nilai K jumlah cluster

Jumlah cluster sebanyak 2 cluster, cluster yang dibentuk yaitu cluster minimum (C1) dan cluster maximum (C2).

3. Menentukan nilai centroid (pusat cluster)

Penentu pusat cluster awal ditentukan secara random yang diambil dari data yang didalam range. Adapun nilai untuk cluster minimum (C1) di ambil dari nilai terendah yang diambil dari tabel 1 dan untuk nilai cluster maximum diambil dari nilai tertinggi pada tabel 1. berikut daftar centroid data acak yang dapat dilihat pada tabel 4

Tabel 4. Menentukan Nilai Centroid Awal

Cluster	Provinsi	Tahun 2020	Tahun 2021	Jumlah
Cluster 1	Gorontalo	3.368	8.462	11.830
Cluster 2	DKI Jakarta	163.111	638.169	801.280

Jumlah Terendah = 11.830

Jumlah Tertinggi = 801.280

4. Menghitung jarak setiap data terhadap centroid (pusat cluster)

Setelah data nilai pusat cluster awal ditentukan. Langkah selanjutnya adalah menghitung jarak masing-masing data terhadap titik pusat cluster. Proses pencarian jarak terpendek pada iterasi 1 dapat dilihat pada perhitungan dan tabel dibawah ini :

Iterasi 1

Perhitungan Provinsi Aceh

$$C_1 = \sqrt{(8.573 - 3.368)^2 + (29.737 - 8.462)^2} = 21902,45762$$

$$C_1 = \sqrt{(8.573 - 163.111)^2 + (29.737 - 638.169)^2} = 627751,1386$$

Perhitungan Provinsi Sumatera Utara

$$C_1 = \sqrt{(17.632 - 3.368)^2 + (88.203 - 8.462)^2} = 81005,13629$$

$$C_1 = \sqrt{(17.632 - 163.111)^2 + (88.203 - 638.169)^2} = 568884.311$$

Perhitungan Provinsi Sumatera Barat

$$C_1 = \sqrt{(22.458 - 3.368)^2 + (67.310 - 8.462)^2} = 61866,91526$$

$$C_1 = \sqrt{(22.458 - 163.111)^2 + (67.310 - 638.169)^2} = 587931,3432$$

Perhitungan Provinsi Riau

$$C_1 = \sqrt{(23.727 - 3.368)^2 + (104.718 - 8.462)^2} = 98385,49902$$

$$C_1 = \sqrt{(23.727 - 163.111)^2 + (104.718 - 638.169)^2} = 551360,0175$$

Perhitungan Provinsi Jambi

$$C_1 = \sqrt{(2.936 - 3.368)^2 + (26.805 - 8.462)^2} = 18348,08636$$

$$C_1 = \sqrt{(2.936 - 163.111)^2 + (26.805 - 638.169)^2} = 631998,3949$$

Hasil dari keseluruhan perhitungan diatas dapat dilihat pada tabel 5 dibawah ini

Tabel 5. Hasil Perhitungan Jarak Pusat Iterasi 1

No	Provinsi	Tahun 2020	Tahun 2021	C1	C2	Jarak Terpendek
1	Aceh	8.573	29.737	21902,45762	627751,1386	21902,45762
2	Sumatra Utara	17.623	88.203	81005,13629	568884,311	81005,13629
3	Sumatra Barat	22.458	67.310	61866,91526	587931,3432	61866,91526
4	Riau	23.727	104.718	98385,49902	551360,0175	98385,49902
5	Jambi	2.936	26.805	18348,08636	631998,3949	18348,08636
6	Sumatra Selatan	10.847	49.019	41240,82553	608508,0511	41240,82553
7	Bengkulu	3.076	20.015	11556,68953	638533,9215	11556,68953
8	Lampung	5.332	44.231	35822,87896	614537,6804	35822,87896
9	Kepulauan Bangka Belitung	1.697	50.297	41868,35877	609629,3692	41868,35877
10	Kepulauan Riau	6.528	47.316	38982,28977	611249,1337	38982,28977
11	Banten	16.163	116.141	108436,5209	542316,2799	108436,5209
12	DKI Jakarta	163.111	638.169	649652,7779	0	0
13	Jawa Barat	73.946	631.781	627302,0244	89393,53315	89393,53315
14	Jawa Tengah	71.531	413.525	410758,1144	242593,947	242593,947
15	DI Yogyakarta	9.497	146.331	138005,1659	515268,7447	138005,1659
16	Jawa Timur	75.274	322.930	322584,2461	327247,5603	322584,2461
17	Bali	16.176	97.668	90120,78173	560117,1531	90120,78173
18	Nusa Tenggara Barat	5.238	22.743	14402,91155	635352,6939	14402,91155
19	Nusa Tenggara Timur	1.808	61.807	53367,80514	598507,989	53367,80514
20	Kalimantan Barat	2.950	38.150	29690,94252	621026,8483	29690,94252
21	Kalimantan Tengah	8.617	36.919	28937,04978	620781,7318	28937,04978
22	Kalimantan Selatan	14.586	55.236	48100,42204	601556,7788	48100,42204
23	Kalimantan Timur	24.081	133.843	127080,3822	523138,6577	127080,3822
24	Kalimantan Utara	2.806	32.946	24490,44916	626093,1023	24490,44916
25	Sulawesi Utara	8.937	25.678	18094,32002	631597,0609	18094,32002

26	Sulawesi Tengah	2.773	44.229	35771,9487	615201,5912	35771,9487
27	Sulawesi Selatan	25.763	83.907	78698,69154	571026,127	78698,69154
28	Sulawesi Tenggara	7.550	12.573	5864,251444	644646,8645	5864,251444
29	Gorontalo	3.368	8.462	0	649652,7779	0
30	Sulawesi Barat	1.707	10.615	2719,251735	647977,8346	2719,251735
31	Maluku	5.448	9.122	2182,200724	648504,2404	2182,200724
32	Maluku Utara	2.654	9.424	1198,015025	648896,5433	1198,015025
33	Papua Barat	5.748	17.341	9192,444778	640461,1771	9192,444778
34	Papua	12.759	21.433	16013,67297	634798,4071	16013,67297

5. Menentukan posisi cluster atau pengelompokan

Dalam menentukan sebuah posisi cluster masing-masing pada data kasus positif COVID-19 penulis melihat berdasarkan jarak terpendek pada C1 dan C2, setiap data yang akan menjadi pusat cluster, dapat dilihat pada tabel 6 dibawah ini.

Tabel 6. Hasil Cluster Iterasi ke-1

No	Provinsi	Tahun 2020	Tahun 2021	C1	C2
1	Aceh	8.573	29.737	1	
2	Sumatra Utara	17.623	88.203	1	
3	Sumatra Barat	22.458	67.310	1	
4	Riau	23.727	104.718	1	
5	Jambi	2.936	26.805	1	
6	Sumatra Selatan	10.847	49.019	1	
7	Bengkulu	3.076	20.015	1	
8	Lampung	5.332	44.231	1	
9	Kepulauan Bangka Belitung	1.697	50.297	1	
10	Kepulauan Riau	6.528	47.316	1	
11	Banten	16.163	116.141	1	
12	DKI Jakarta	163.111	638.169		1
13	Jawa Barat	73.946	631.781		1
14	Jawa Tengah	71.531	413.525		1
15	DI Yogyakarta	9.497	146.331	1	
16	Jawa Timur	75.274	322.930	1	
17	Bali	16.176	97.668	1	
18	Nusa Tenggara Barat	5.238	22.743	1	
19	Nusa Tenggara Timur	1.808	61.807	1	
20	Kalimantan Barat	2.950	38.150	1	
21	Kalimantan Tengah	8.617	36.919	1	
22	Kalimantan Selatan	14.586	55.236	1	
23	Kalimantan Timur	24.081	133.843	1	
24	Kalimantan Utara	2.806	32.946	1	
25	Sulawesi Utara	8.937	25.678	1	
26	Sulawesi Tengah	2.773	44.229	1	
27	Sulawesi Selatan	25.763	83.907	1	
28	Sulawesi Tenggara	7.550	12.573	1	
29	Gorontalo	3.368	8.462	1	
30	Sulawesi Barat	1.707	10.615	1	
31	Maluku	5.448	9.122	1	
32	Maluku Utara	2.654	9.424	1	
33	Papua Barat	5.748	17.341	1	
34	Papua	12.759	21.433	1	

Proses *K-means* akan terus berinteraksi sampai pengelompokan data sama dengan pengelompokan data dengan interaksi sebelumnya. Sebuah proses data akan terus melakukan interaksi sampai data pada interaksi terakhir sama dengan data interaksi sebelumnya.

6. Menghitung centroid baru

Dengan mrnggunakan hasil dari setiap anggota masing-masing cluster, setelah didapatkan hasil jarak di setiap objek pada iterasi ke-1 maka lanjut iterasi ke-2 pada perhitungan dibawah ini :

C1 Tahun 2020

$$C1 = (8.573 + 17.623 + 22.458 + 23.727 + 2.936 + 10.847 + 3.076 + 5.332 + 1.697 + 6.528 + 16.163 + 9.497 + 75.274 + 16.176 + 5.238 + 1.808 + 2.950 + 8.617 + 14.586 + 24.081 + 2.806 + 8.937 + 2.773 + 25.763 + 7.550 + 3.368 + 1.707 + 5.448 + 2.654 + 5.748 + 12.759) / 31 = 11.506$$

C1 Tahun 2021

$$C1 = (29.737 + 88.205 + 67.310 + 104.718 + 26.805 + 49.019 + 20.015 + 44.231 + 50.297 + 47.316 + 116.141 + 146.331 + 322.930 + 97.668 + 22.743 + 61.807 + 38.150 + 36.919 + 55.236 + 133.843 + 32.946 + 25.678 + 44.229 + 83.907 + 12.573 + 8.462 + 10.615 + 9.122 + 9.424 + 17.341 + 21.433) / 31 = 59.198$$

C2 Tahun 2020

$$C2 = (163.111 + 73.946 + 71.531) / 3 = 102.863$$

C2 Tahun 2021

$$C2 = (638.169 + 631.781 + 413.525) / 3 = 561.158$$

Hasil dari keseluruhan perhitungan diatas dapat dilihat pada tabel 7 dibawah ini.

Tabel 7. Centroid Baru Iterasi-1

Cluster	Provinsi	Tahun 2020	Tahun 2021
Cluster 1	Gorontalo	11.506	59.198
Cluster 2	DKI Jakarta	102.863	561.158

Pada tabel 7 diatas merupakan hasil centroid baru iterasi ke-1. Selanjutnya dilakukan kembali langkah ke-4 sampai 6, maka jika nilai centroid hasil iterasi dengan nilai centroid sebelumnya bernilai sama serta posisi cluster data tidak mengalami perubahan, maka proses iterasi berhenti. Namun, pada uji coba iterasi serta posisi cluster diatas masih mengalami perubahan, maka proses iterasi berlanjut pada iterasi selanjutnya. Berikut hasil cluster iterasi 1 dapat dilihat pada tabel 8 dibawah ini.

Tabel 8. Centroid Baru Iterasi ke-2

Cluster	Provinsi	Tahun 2020	Tahun 2021
Cluster 1	Gorontalo	11.506	59.198
Cluster 2	DKI Jakarta	102.863	561.158

Langkah selanjutnya adalah melakukan perhitungan jarak terhadap data kasus positif COVID-19 berdasarkan Provinsi Indonesia menggunakan data *centroid* baru. Proses pencarian jarak terpendek pada iterasi 2 dapat dilihat pada perhitungan dibawah ini :

Iterasi 2

Perhitungan Provinsi Aceh

$$C_1 = \sqrt{(8.573 - 11.506)^2 + (29.737 - 59.196)^2} = 29607,03577$$

$$C_2 = \sqrt{(8.573 - 102.863)^2 + (29.737 - 561.158)^2} = 539721,3862$$

Perhitungan Provinsi Sumatera Utara

$$C_1 = \sqrt{(17.623 - 11.506)^2 + (88.205 - 59.196)^2} = 29642,56408$$

$$C_2 = \sqrt{(17.632 - 102.863)^2 + (88.205 - 561.158)^2} = 480575,2263$$

Perhitungan Provinsi Sumatera Barat

$$C_1 = \sqrt{(22.458 - 11.506)^2 + (67.310 - 59.196)^2} = 13628,47017$$

$$C_2 = \sqrt{(22.458 - 102.863)^2 + (67.310 - 561.158)^2} = 500350,9636$$

Perhitungan Provinsi Riau

$$C_1 = \sqrt{(23.727 - 11.506)^2 + (104.718 - 59.196)^2} = 47131,51704$$

$$C_2 = \sqrt{(23.727 - 102.863)^2 + (104.718 - 561.158)^2} = 463249,6429$$

Perhitungan Provinsi Jambi

$$C_1 = \sqrt{(2.936 - 11.506)^2 + (26.805 - 59.196)^2} = 33507,94053$$

$$C_2 = \sqrt{(2.936 - 102.863)^2 + (26.805 - 561.158)^2} = 543616,4305$$

Melakukan perhitungan jarak terhadap data-data kasus positif COVID-19 berdasarkan Jumlah kasus Perprovinsi Indonesia dengan menggunakan *centroid* baru. Hasil dari keseluruhan perhitungan dapat dilihat pada tabel 9 sebagai berikut:

Tabel 9. Hasil Perhitungan Jarak Pusat Iterasi 2

No	Provinsi	Tahun 2020	Tahun 2021	C1	C2	Jarak Terpendek
1	Aceh	8.573	29.737	29607,03577	539721,3862	29607,03577
2	Sumatra Utara	17.623	88.203	29642,56408	480575,2263	29642,56408
3	Sumatra Barat	22.458	67.310	13628,47017	500350,9636	13628,47017
4	Riau	23.727	104.718	47131,51704	463249,6429	47131,51704
5	Jambi	2.936	26.805	33507,94053	543616,4305	33507,94053
6	Sumatra Selatan	10.847	49.019	10200,69318	520339,8694	10200,69318
7	Bengkulu	3.076	20.015	40080,0176	550266,7408	40080,0176
8	Lampung	5.332	44.231	16190,90991	526047,6204	16190,90991
9	Kepulauan Bangka Belitung	1.697	50.297	13246,11109	520781,9064	13246,11109
10	Kepulauan Riau	6.528	47.316	12883,14158	522794,7126	12883,14158
11	Banten	16.163	116.141	57132,72513	453384,2291	57132,72513
12	DKI Jakarta	163.111	638.169	598490,5572	97777,8321	97777,8321
13	Jawa Barat	73.946	631.781	575977,0679	76313,39763	76313,39763
14	Jawa Tengah	71.531	413.525	359374,8988	150921,4181	150921,4181
15	DI Yogyakarta	9.497	146.331	87155,81305	425204,4969	87155,81305
16	Jawa Timur	75.274	322.930	271331,312	239820,5023	239820,5023
17	Bali	16.176	97.668	38752,01002	471527,1649	38752,01002
18	Nusa Tenggara Barat	5.238	22.743	36990,35526	547194,3409	36990,35526
19	Nusa Tenggara Timur	1.808	61.807	10043,15654	509474,0423	10043,15654
20	Kalimantan Barat	2.950	38.150	22721,04984	532466,2033	22721,04984
21	Kalimantan Tengah	8.617	36.919	22465,94273	532643,5246	22465,94273
22	Kalimantan Selatan	14.586	55.236	5018,353728	513566,1372	5018,353728
23	Kalimantan Timur	24.081	133.843	75696,38247	434516,91	75696,38247
24	Kalimantan Utara	2.806	32.946	27656,53617	537605,4368	27656,53617
25	Sulawesi Utara	8.937	25.678	33618,6893	543655,4223	33618,6893
26	Sulawesi Tengah	2.773	44.229	17330,74613	526530,0343	17330,74613
27	Sulawesi Selatan	25.763	83.907	28526,58966	483438,9245	28526,58966
28	Sulawesi Tenggara	7.550	12.573	46792,91852	556803,7108	46792,91852
29	Gorontalo	3.368	8.462	51384,94037	561580,2931	51384,94037
30	Sulawesi Barat	1.707	10.615	49561,79597	559759,2614	49561,79597
31	Maluku	5.448	9.122	50441,51217	560565,5453	50441,51217
32	Maluku Utara	2.654	9.424	50555,43788	560760,69	50555,43788
33	Papua Barat	5.748	17.341	42251,60256	552420,6282	42251,60256
34	Papua	12.759	21.433	37786,12051	547194,7607	37786,12051

Terdapat hasil cluster ke-2 pada pengelompokan iterasi ke-2, dapat dilihat pada tabel 10 dibawah ini .

Tabel 10. Hasil Cluster ke-2

No	Provinsi	Tahun 2020	Tahun 2021	C1	C2
1	Aceh	8.573	29.737	1	
2	Sumatra Utara	17.623	88.203	1	
3	Sumatra Barat	22.458	67.310	1	
4	Riau	23.727	104.718	1	
5	Jambi	2.936	26.805	1	
6	Sumatra Selatan	10.847	49.019	1	
7	Bengkulu	3.076	20.015	1	
8	Lampung	5.332	44.231	1	
9	Kepulauan Bangka Belitung	1.697	50.297	1	

No	Provinsi	Tahun 2020	Tahun 2021	C1	C2
10	Kepulauan Riau	6.528	47.316	1	
11	Banten	16.163	116.141	1	
12	DKI Jakarta	163.111	638.169		1
13	Jawa Barat	73.946	631.781		1
14	Jawa Tengah	71.531	413.525		1
15	DI Yogyakarta	9.497	146.331	1	
16	Jawa Timur	75.274	322.930		1
17	Bali	16.176	97.668	1	
18	Nusa Tenggara Barat	5.238	22.743	1	
19	Nusa Tenggara Timur	1.808	61.807	1	
20	Kalimantan Barat	2.950	38.150	1	
21	Kalimantan Tengah	8.617	36.919	1	
22	Kalimantan Selatan	14.586	55.236	1	
23	Kalimantan Timur	24.081	133.843	1	
24	Kalimantan Utara	2.806	32.946	1	
25	Sulawesi Utara	8.937	25.678	1	
26	Sulawesi Tengah	2.773	44.229	1	
27	Sulawesi Selatan	25.763	83.907	1	
28	Sulawesi Tenggara	7.550	12.573	1	
29	Gorontalo	3.368	8.462	1	
30	Sulawesi Barat	1.707	10.615	1	
31	Maluku	5.448	9.122	1	
32	Maluku Utara	2.654	9.424	1	
33	Papua Barat	5.748	17.341	1	
34	Papua	12.759	21.433	1	

Menghitung *centroid* baru dengan menggunakan hasil dari setiap anggota pada masing-masing *cluster*, setelah didapatkan hasil jarak di setiap objek pada iterasi ke 2 maka lanjut ke iterasi ke 3 pada perhitungan dibawah ini :

C1 Tahun 2020

$$C1 = (8.573 + 17.623 + 22.458 + 23.727 + 2.936 + 10.847 + 3.076 + 5.332 + 1.697 + 6.528 + 16.163 + 9.497 + 16.176 + 5.238 + 1.808 + 2.950 + 8.617 + 14.586 + 24.081 + 2.806 + 8.937 + 2.773 + 25.763 + 7.550 + 3.368 + 1.707 + 5.448 + 2.654 + 5.748 + 12.759) / 30 = 9.381$$

C1 Tahun 2021

$$C1 = (29.737 + 88.205 + 67.310 + 104.718 + 26.805 + 49.019 + 20.015 + 44.231 + 50.297 + 47.316 + 116.141 + 146.331 + 97.668 + 22.743 + 61.807 + 38.150 + 36.919 + 55.236 + 133.843 + 32.946 + 25.678 + 44.229 + 83.907 + 12.573 + 8.462 + 10.615 + 9.122 + 9.424 + 17.341 + 21.433) / 30 = 50.407$$

C2 Tahun 2020

$$C2 = (163.111 + 73.946 + 71.531 + 75.274) / 4 = 95.966$$

C2 Tahun 2021

$$C2 = (638.169 + 631.781 + 413.525 + 322.930) / 4 = 501.601$$

Hasil dari keseluruhan perhitungan diatas dapat dilihat pada tabel 11 dibawah ini .

Tabel 11. Centroid Baru Iterasi ke-2

Cluster	Provinsi	Tahun 2020	Tahun 2021
Cluster 1	Gorontalo	9.381	50.407
Cluster 2	DKI Jakarta	95.966	501.601

Pada tabel 9 diatas merupakan hasil *centroid* baru iterasi ke 2. Selanjutnya dilakukan kembali langkah ke 4 sampai 6 . maka jika nilai *centroid* hasil iterasi dengan nilai *centroid* sebelumnya bernilai sama serta posisi *cluster* data tidak mengalami perubahan maka proses iterasi berhenti. Namun pada uji coba iterasi serta posisi *cluster* diatas masih mengalami perubahan maka proses iterasi berlanjut pada iterasi selanjutnya. Berikut tabel *centroid* baru untuk iterasi ke 3, dapat dilihat pada tabel 12 dibawah ini :

Tabel 12. Centroid Baru Untuk Iterasi ke-3

Cluster	Provinsi	Tahun 2020	Tahun 2021
Cluster 1	Gorontalo	9.381	50.407
Cluster 2	DKI Jakarta	95.966	501.601

Langkah selanjutnya adalah melakukan perhitungan jarak terhadap data kasus positif COVID-19 berdasarkan Provinsi Indonesia menggunakan data centroid baru. Proses pencarian jarak terpendek pada iterasi 3 dapat dilihat pada perhitungan dan tabel dibawah ini.

Iterasi ke 3

Perhitungan Provinsi Aceh

$$C_1 = \sqrt{(8.573 - 9.381)^2 + (29.737 - 50.047)^2} = 20686,08108$$

$$C_2 = \sqrt{(8.573 - 95.966)^2 + (29.737 - 501.601)^2} = 479888,8616$$

Perhitungan Provinsi Sumatera Utara

$$C_1 = \sqrt{(17.623 - 9.381)^2 + (88.203 - 50.047)^2} = 38683,94629$$

$$C_2 = \sqrt{(17.623 - 95.966)^2 + (88.203 - 501.601)^2} = 420756,0581$$

Perhitungan Provinsi Sumatera Barat

$$C_1 = \sqrt{(22.458 - 9.381)^2 + (67.310 - 50.047)^2} = 21370,83722$$

$$C_2 = \sqrt{(22.458 - 95.966)^2 + (67.310 - 501.601)^2} = 440468,2081$$

Perhitungan Provinsi Riau

$$C_1 = \sqrt{(23.727 - 9.381)^2 + (104.718 - 50.047)^2} = 56173,51401$$

$$C_2 = \sqrt{(23.727 - 95.966)^2 + (104.718 - 501.601)^2} = 403403,9105$$

Perhitungan Provinsi Jambi

$$C_1 = \sqrt{(2.936 - 9.381)^2 + (26.805 - 50.047)^2} = 24466,40292$$

$$C_2 = \sqrt{(2.936 - 95.966)^2 + (26.805 - 501.601)^2} = 483824,3141$$

Melakukan perhitungan jarak terhadap data-data kasus positif COVID-19 berdasarkan Jumlah kasus Perprovinsi Indonesia dengan menggunakan *centroid* baru. Hasil dari keseluruhan perhitungan dapat dilihat pada tabel 13 sebagai berikut:

Tabel 13. Hasil Perhitungan Jarak Terdekat Pusat Iterasi 3

No	Provinsi	Tahun 2020	Tahun 2021	C1	C2	Jarak Terpendek
1	Aceh	8.573	29.737	20686,08108	479888,8616	20686,08108
2	Sumatra Utara	17.623	88.203	38683,94629	420756,0581	38683,94629
3	Sumatra Barat	22.458	67.310	21370,83722	440468,2081	21370,83722
4	Riau	23.727	104.718	56173,51401	403403,9105	56173,51401
5	Jambi	2.936	26.805	24466,40292	483824,3141	24466,40292
6	Sumatra Selatan	10.847	49.019	2019,139381	460516,94	2019,139381
7	Bengkulu	3.076	20.015	31039,38213	490462,8176	31039,38213
8	Lampung	5.332	44.231	7385,120376	466263,849	7385,120376
9	Kepulauan Bangka Belitung	1.697	50.297	7684,65829	461044,549	7684,65829
10	Kepulauan Riau	6.528	47.316	4206,54061	463005,5667	4206,54061
11	Banten	16.163	116.141	66082,65013	393634,4031	66082,65013
12	DKI Jakarta	163.111	638.169	607533,3488	152181,6957	152181,6957
13	Jawa Barat	73.946	631.781	584947,8913	132028,882	132028,882
14	Jawa Tengah	71.531	413.525	368398,0227	91402,7932	91402,7932
15	DI Yogyakarta	9.497	146.331	95923,7703	365641,5622	95923,7703
16	Jawa Timur	75.274	322.930	280375,6891	179865,3767	179865,3767
17	Bali	16.176	97.668	47746,70252	411738,3086	47746,70252
18	Nusa Tenggara Barat	5.238	22.743	27972,78747	487377,3721	27972,78747
19	Nusa Tenggara Timur	1.808	61.807	13685,81271	449760,622	13685,81271
20	Kalimantan Barat	2.950	38.150	13841,87304	472693,2878	13841,87304
21	Kalimantan Tengah	8.617	36.919	13509,91226	472820,6361	13509,91226
22	Kalimantan Selatan	14.586	55.236	7099,982867	453722,9986	7099,982867
23	Kalimantan Timur	24.081	133.843	84720,7764	374717,9096	84720,7764

No	Provinsi	Tahun 2020	Tahun 2021	C1	C2	Jarak Terpendek
24	Kalimantan Utara	2.806	32.946	18658,13145	477824,6915	18658,13145
25	Sulawesi Utara	8.937	25.678	24733,28316	483814,9437	24733,28316
26	Sulawesi Tengah	2.773	44.229	9046,286132	466769,9831	9046,286132
27	Sulawesi Selatan	25.763	83.907	37290,80574	423552,6856	37290,80574
28	Sulawesi Tenggara	7.550	12.573	37878,57348	496956,6681	37878,57348
29	Gorontalo	3.368	8.462	42374,0812	501757,528	42374,0812
30	Sulawesi Barat	1.707	10.615	40525,49036	499952,1602	40525,49036
31	Maluku	5.448	9.122	41472,20077	500728,6985	41472,20077
32	Maluku Utara	2.654	9.424	41531,69409	500944,5892	41531,69409
33	Papua Barat	5.748	17.341	33265,26591	492592,3132	33265,26591
34	Papua	12.759	21.433	29170,56471	487324,1939	29170,56471

Terdapat hasil cluster ke-3 pada pengelompokan iterasi 3, dapat dilihat pada tabel 14 dibawah ini :

Tabel 14. Hasil Cluster ke 3

No	Provinsi	Tahun 2020	Tahun 2021	C1	C2
1	Aceh	8.573	29.737	1	
2	Sumatra Utara	17.623	88.203	1	
3	Sumatra Barat	22.458	67.310	1	
4	Riau	23.727	104.718	1	
5	Jambi	2.936	26.805	1	
6	Sumatra Selatan	10.847	49.019	1	
7	Bengkulu	3.076	20.015	1	
8	Lampung	5.332	44.231	1	
9	Kepulauan Bangka Belitung	1.697	50.297	1	
10	Kepulauan Riau	6.528	47.316	1	
11	Banten	16.163	116.141	1	
12	DKI Jakarta	163.111	638.169	1	
13	Jawa Barat	73.946	631.781	1	
14	Jawa Tengah	71.531	413.525	1	
15	DI Yogyakarta	9.497	146.331	1	
16	Jawa Timur	75.274	322.930	1	
17	Bali	16.176	97.668	1	
18	Nusa Tenggara Barat	5.238	22.743	1	
19	Nusa Tenggara Timur	1.808	61.807	1	
20	Kalimantan Barat	2.950	38.150	1	
21	Kalimantan Tengah	8.617	36.919	1	
22	Kalimantan Selatan	14.586	55.236	1	
23	Kalimantan Timur	24.081	133.843	1	
24	Kalimantan Utara	2.806	32.946	1	
25	Sulawesi Utara	8.937	25.678	1	
26	Sulawesi Tengah	2.773	44.229	1	
27	Sulawesi Selatan	25.763	83.907	1	
28	Sulawesi Tenggara	7.550	12.573	1	
29	Gorontalo	3.368	8.462	1	
30	Sulawesi Barat	1.707	10.615	1	
31	Maluku	5.448	9.122	1	
32	Maluku Utara	2.654	9.424	1	
33	Papua Barat	5.748	17.341	1	
34	Papua	12.759	21.433	1	

Menghitung *centroid* baru dengan menggunakan hasil dari setiap anggota pada masing-masing *cluster*, setelah didapatkan hasil jarak di setiap objek pada iterasi ke 3 maka lanjut ke iterasi ke 4 pada perhitungan dibawah ini :

C1 Tahun 2020

$$C1 = (8.573 + 17.623 + 22.458 + 23.727 + 2.936 + 10.847 + 3.076 + 5.332 + 1.697 + 6.528 + 16.163 + 9.497 + 16.176 + 5.238 + 1.808 + 2.950 + 8.617 + 14.586 + 24.081 + 2.806 + 8.937 + 2.773 + 25.763 + 7.550 + 3.368 + 1.707 + 5.448 + 2.654 + 5.748 + 12.759) / 30 = 9.381$$

C1 Tahun 2021

$$C1 = (29.737 + 88.205 + 67.310 + 104.718 + 26.805 + 49.019 + 20.015 + 44.231 + 50.297 + 47.316 + 116.141 + 146.331 + 97.668 + 22.743 + 61.807 + 38.150 + 36.919 + 55.236 + 133.843 + 32.946 + 25.678 + 44.229 + 83.907 + 12.573 + 8.462 + 10.615 + 9.122 + 9.424 + 17.341 + 21.433) / 30 = 50.407$$

C2 Tahun 2020

$$C2 = (163.111 + 73.946 + 71.531 + 75.274) / 4 = 95.966$$

C2 Tahun 2021

$$C2 = (638.169 + 631.781 + 413.525 + 322.930) / 4 = 501.601$$

Dari perhitungan diatas menghasilkan centroid baru dari interaksi ke 3 , dapat dilihat pada tabel 15 dibawah ini :

Tabel 15. Centroid Baru Iterasi ke 3

Cluster	Provinsi	Tahun 2020	Tahun 2021
Cluster 1	Gorontalo	9.381	50.407
Cluster 2	DKI Jakarta	95.966	501.601

Hasil perhitungan manual mendapatkan hasil akhir yang dimana pada iterasi ke 2 dan iterasi ke 3 pengelompokkan data yang dilakukan terhadap 2 cluster dengan hasil yang sama. Hasil dari kedua cluster tersebut bernilai C1= 30 dan C2= 4 dan pada posisi data tiap *cluster* sama. Sehingga posisi *cluster* pada data tersebut tidak mengalami perubahan lagi diantara iterasi ke 2 dan iterasi ke 3 maka proses iterasi berhenti sampai iterasi 3.

4. KESIMPULAN

Hasil dari penelitian yang diperoleh dapat disimpulkan bahwa perhitungan manual pada microsoft excel menghasilkan hasil yang sama dengan yang dikombinasikan oleh software rapidminer versi 5.3 bahwa metode tersebut mampu menyelesaikan dari sebuah pengelompokkan berdasarkan kasus positif COVID-19 memiliki hasil yang akurat. Hasil yang diperoleh dari metode k-means clustering yang di implementasikan ke dalam rapidminer memiliki nilai validasi yang sama yaitu menghasilkan beberapa cluster yaitu cluster terendah C1= 30 items diantaranya, Aceh, Sumatra Utara, Sumatra Barat, Riau, Jambi, Sumatra Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, Banten, Daerah Istimewah Yogyakarta, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, Papua. 65 Sedangkan untuk cluster tertinggi memiliki C2= 4 items diantaranya DKI Jakarta, Jawa Tengah, Jawa Barat, Jawa Timur.

REFERENCES

- [1] S. N. Alaziz, A. A. Alshowiman, B. Albayati, A. al A. H. El-Bagoury, and W. Shafik, "Clustering of COVID-19 Multi-Time Series-Based K-Means and PCA With Forecasting," *Int. J. Data Warehous. Min.*, vol. 19, no. 3, 2023, doi: 10.4018/IJDWM.317374.
- [2] D. T. Utari, "Analisis Karakteristik Wilayah Transmisi Covid-19 dengan Menggunakan Metode K-Means Clustering," *J. Media Tek. dan Sist. Ind.*, vol. 5, no. 1, p. 25, 2021, doi: 10.35194/jmtsi.v5i1.1220.
- [3] I. S. Mangku Negara, P. Purwono, and I. A. Ashari, "Analisa Cluster Data Transaksi Penjualan Minimarket Selama Pandemi Covid-19 dengan Algoritma K-means," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 6, no. 3, p. 153, 2021, doi: 10.31328/jointecs.v6i3.2693.
- [4] U. R. Gurning and M. Mustakim, "Penerapan Algoritma K-Means dan K-Medoid untuk Pengelompokkan Data Pasien Covid-19," *Build. Informatics, Technol. Sci.*, vol. 3, no. 1, pp. 48–55, 2021, doi: 10.47065/bits.v3i1.1003.
- [5] S. K. Sunori, P. B. Negi, S. Maurya, P. Juneja, A. Rana, and Bhawana, "K-Means Clustering of Ambient Air Quality Data of Uttarakhand, India during Lockdown Period of Covid-19 Pandemic," *Proc. 6th Int. Conf. Inven. Comput. Technol. ICICT 2021*, pp. 1254–1259, 2021, doi: 10.1109/ICICT50816.2021.9358627.
- [6] D. D. Darmansah, "Analisis Penyebaran Penularan Virus Covid-19 di Provinsi Jawa Barat Menggunakan Algoritma K-Means Clustering," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 3, pp. 1188–1199, 2021, doi: 10.35957/jatisi.v8i3.1034.
- [7] A. L. R. Putri and N. Dwidayati, "Analisa Perbandingan K-Means dan Fuzzy C-Menas dalam Pengelompokan Daerah Penyebaran Covid-19 Indonesia," *UNNES J. Math.*, vol. 10, no. 2, pp. 50–55, 2021, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/ujm>.
- [8] H. Gunawan and V. Purwayoga, "Data Mining Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Potensi Penyebaran Virus Corona Di Kota Cirebon," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 11, no. 1, pp. 1–8, 2022, doi: 10.32736/sisfokom.v11i1.1316.
- [9] N. G. Ali, S. D. Abed, F. A. J. Shaban, K. Tongkachok, S. Ray, and R. A. Jaleel, "Hybrid of K-Means and partitioning around medoids for predicting COVID-19 cases: Iraq case study," *Period. Eng. Nat. Sci.*, vol. 9, no. 4, pp. 569–579, 2021, doi: 10.21533/pen.v9i4.2382.
- [10] V. Chandu, "Identification of spatial variations in COVID-19 epidemiological data using K-Means clustering algorithm: a global perspective," *medRxiv*, p. 2020.06.03.20121194, 2020, [Online]. Available: <http://medrxiv.org/content/early/2020/06/05/2020.06.03.20121194.abstract>.
- [11] W. Utomo, "The comparison of k-means and k-medoids algorithms for clustering the spread of the covid-19 outbreak in

- Indonesia,” *Ilk. J. Ilm.*, vol. 13, no. 1, pp. 31–35, 2021, doi: 10.33096/ilkom.v13i1.763.31-35.
- [12] Darmansah and N. W. Wardani, “Analisis Pesebaran Penularan Virus Corona di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 105–117, 2021, doi: 10.35957/jatisi.v8i1.590.
- [13] A. Mahmudan, “Clustering of District or City in Central Java Based COVID-19 Case Using K-Means Clustering,” *J. Mat. Stat. dan Komputasi*, vol. 17, no. 1, pp. 1–13, 2020, doi: 10.20956/jmsk.v17i1.10727.
- [14] M. Zubair, M. Asif Iqbal, A. Shil, E. Haque, M. Moshikul Hoque, and I. H. Sarker, “An Efficient K-Means Clustering Algorithm for Analysing COVID-19,” *Adv. Intell. Syst. Comput.*, vol. 1375 AIST, pp. 422–432, 2021, doi: 10.1007/978-3-030-73050-5_43.
- [15] N. Mirantika, “Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Covid-19 di Provinsi Jawa Barat,” *Nuansa Inform.*, vol. 15, no. 2, pp. 92–98, 2021, doi: 10.25134/nuansa.v15i2.4321.
- [16] R. Adha, N. Nurhaliza, U. Sholeha, and M. Mustakim, “Perbandingan Algoritma DBSCAN dan K-Means Clustering untuk Pengelompokan Kasus Covid-19 di Dunia,” *SITEKIN J. Sains, Teknol. dan Ind.*, vol. 18, no. 2, pp. 206–211, 2021.
- [17] F. Virgantari and Y. E. Faridhan, “K-Means Clustering of COVID-19 Cases in Indonesia’s Provinces,” 2020.
- [18] J. Hutagalung, N. L. W. S. R. Ginantra, G. W. Bhawika, W. G. S. Parwita, A. Wanto, and P. D. Panjaitan, “COVID-19 Cases and Deaths in Southeast Asia Clustering using K-Means Algorithm,” *J. Phys. Conf. Ser.*, vol. 1783, no. 1, 2021, doi: 10.1088/1742-6596/1783/1/012027.
- [19] T. Hardiani, “Analisis Clustering Kasus Covid 19 di Indonesia Menggunakan Algoritma K-Means,” *J. Nas. Pendidik. Tek. Inform.*, vol. 11, no. 2, pp. 156–165, 2022, doi: 10.23887/janapati.v11i2.45376.
- [20] Z. Nabila, A. Rahman Isnain, and Z. Abidin, “Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means,” *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSL>.