

## Penerapan Algoritma K-Means Untuk Mengelompokkan Provinsi Di Indonesia Berdasarkan Data Sebaran Covid-19

Achmad Furqon Nur Fitriadhi<sup>a</sup>, Retno Wahyusari<sup>b\*</sup>

<sup>ab</sup>Sekolah Tinggi Teknologi Ronggolawe

\*E-mail: [retnowahyusari@gmail.com](mailto:retnowahyusari@gmail.com)

Tentang naskah:

-diterima, 27 Des 2021

-direview, 12 Jan 2022

-diterbitkan, 21 Jan 2022

### Intisari

Pada bulan Desember 2019 di Wuhan, Cina terjadi wabah yang menyerang saluran pernafasan. Awal bulan Januari, WHO mengidentifikasi virus tersebut sebagai Coronavirus atau 2019-nCoV yang kemudian mengumumkan nama resmi untuk virus yang melanda adalah COVID-19. Virus COVID-19 tidak hanya terjadi di Wuhan, tetapi juga menyebar ke seluruh dunia. Hal ini tidak terkecuali di Indonesia, sejak Maret hingga April data grafik semakin meningkat signifikan. Tujuan dari penelitian ini adalah menerapkan algoritma *K-Means* dalam mengelompokkan provinsi berdasarkan tingkat persebaran virus Corona (COVID-19). Penelitian yang dilakukan memiliki keterbaruan dengan menambahkan atribut dalam penentuan kelompok, atribut yang digunakan adalah jumlah positif, jumlah sembuh dan jumlah meninggal. Penelitian menghasilkan 2 (dua) kelompok dengan keanggotaan kluster 0 menunjukkan wilayah yang paling sedikit terdampak covid-19 dengan jumlah 6 data, dan cluster 1 menunjukkan wilayah yang paling banyak terdampak covid-19 dengan jumlah 28 data. Dari penelitian ini maka dapat disimpulkan bahwa pemerintah sebaiknya melakukan penanganan covid-19 lebih terfokus pada cluster 1 yang jumlah terdampaknya paling banyak.

Kata kunci:

*Algoritma K-Means*,  
Covid-19, Penambangan  
Data

### Abstract

*In December 2019 in Wuhan, China there was an outbreak that attacked the respiratory tract. Early in January, WHO identified the virus as Coronavirus or 2019-nCoV which later announced the official name for the virus that was ravaging COVID-19. The COVID-19 virus did not only occur in Wuhan, but also spread throughout the world. This is no exception in Indonesia, from March to April the graph data has increased significantly. The purpose of this study is to apply the K-Means algorithm in grouping provinces based on the level of spread of the Corona virus (COVID-19). The research carried out is up-to-date by adding attributes in group determination, the attributes used are the number of positives, the number of recoveries and the number of deaths. The study resulted in 2 (two) groups with cluster 0 membership indicating the area that was least affected by COVID-19 with a total of 6 data, and cluster 1 showing the area most affected by COVID-19 with a total of 28 data. From this research, it can be concluded that the government should carry out handling of COVID-19 more focused on cluster 1 which has the largest number of impacts.*

Keyword:

*K-Means Algorithm*,  
Covid-19, Data Mining

### 1. Pendahuluan

Pada bulan Desember 2019 di Wuhan, Cina terjadi wabah yang menyerang saluran pernafasan. Awal bulan Januari, WHO mengidentifikasi virus tersebut sebagai Coronavirus atau 2019-nCoV yang kemudian mengumumkan nama resmi untuk virus yang melanda adalah COVID-19. Virus corona merupakan keluarga besar dari virus yang menyebabkan flu biasa hingga penyakit seperti MERS atau SARS ([https://promkes.kemkes.go.id/informasi-tentang-virus-corona-novel-](https://promkes.kemkes.go.id/informasi-tentang-virus-corona-novel-coronavirus)

[coronavirus](https://promkes.kemkes.go.id/informasi-tentang-virus-corona-novel-coronavirus),” 2020). Virus COVID-19 tidak hanya terjadi di Wuhan, tetapi juga menyebar ke seluruh dunia. Hal ini tidak terkecuali di Indonesia, sejak Maret hingga April data grafik semakin meningkat signifikan di wilayah Sumatera Utara, Bali, Kepulauan Riau, DKI Jakarta, Sulawesi Selatan dan Sulawesi Utara (<https://www.kompas.com/sains/read/2020/05/11/130600623/diumumkan-awal-maret-ahli-virus-corona-masuk-indonesia-dari-januari>,” 2020). Sampai saat ini penyebaran virus telah keseluruh provinsi yang ada di Indonesia, hal ini terlihat dari peta

sebaran kasus per provinsi pada Gambar 1.1 (<https://covid19.go.id/peta-sebaran> 21 november 2020,” 2020).



Gambar 1. Peta Sebaran COVID-19

Begitu cepatnya penyebaran COVID-19 di tiap provinsi, mengharuskan pemerintah untuk merencanakan berbagai strategi pencegahan berdasarkan tingkat penyebaran yang terjadi pada masing-masing daerah. Berdasarkan data yang di peroleh dari penyebaran virus COVID-19 dapat dilakukan pengelompokan (Noviyanto, 2020). Maka penentuan suatu provinsi masuk kedalam salah satu zona dapat juga dilakukan dengan menggali informasi dari data tingkat sebaran COVID-19 per provinsi. Algoritma penentuan zona dapat dilakukan dengan algoritma *clustering*/pengelompokan.

Penelitian yang dilakukan oleh Noviyanto (Noviyanto, 2020) menggunakan algoritma *K-Means* untuk mengelompokkan berdasarkan jumlah kematian di benua Asia. Hasil penelitian adalah *cluster* jumlah kematian penderita Covid-19 kedalam 3 *cluster*. Terdapat 4 negara dengan *cluster* tingkat tinggi yaitu: Turki, Iran, India, dan China dengan *cluster* tingkat sedang sebanyak 4 negara yaitu : Pakistan, Indonesia, Jepang, dan Piliphina dan dengan *cluster* rendah adalah 41 negara lainnya.

Berdasarkan latar belakang masalah diatas, menyatakan bahwa algoritma *K-Means* mampu menghasilkan keanggotan kelompok dengan baik. Diharapkan dengan mengetahui hasil pengelompokan mampu memberika informasi bagi pemerintah untuk menyusun strategi penangan penyebaran COVID-19.

## 2. Kerangka Teori

### 2.1. Data Mining

Menurut Larose (Larose, 2005) *data mining* adalah proses menemukan korelasi, pola, dan tren baru yang bermakna dengan menyaring sejumlah besar data yang disimpan dalam repositori, menggunakan teknologi pengenalan pola secara teknik statistik

dan matematika. Terdapat beberapa istilah lain yang memiliki makna sama dengan data mining yaitu *Knowledge Discovery in Database* (KDD).

### 2.2. Clustering

*Clustering* (pengelompokan data) mempertimbangkan sebuah pendekatan penting untuk mencari kesamaan dalam data dan menempatkan data yang sama ke dalam kelompok-kelompok. *Clustering* membagi kumpulan data ke dalam beberapa kelompok dimana kesamaan dalam sebuah kelompok adalah lebih besar daripada diantara kelompok-kelompok (Xu & Wunsch, 2009).

### 2.3. Algoritma K-Means

Data *Clustering*/ pengelompokan data merupakan salah satu metode *Data Mining* yang bersifat tanpa arahan (*unsupervised*). *K-Means* secara umum dilakukan dengan algoritma dasar sebagai berikut (Xu & Wunsch, 2009):

- Pilih K buah titik *centroid* secara acak,
- Kelompokkan data sehingga terbentuk K buah *cluster* dengan titik *centroid* dari setiap *cluster* merupakan titik *centroid* yang telah dipilih sebelumnya,
- Perbaharui nilai titik *centroid*,

$$d(x_j, c_j) = \sqrt{\sum_{j=1}^n (x_j - c_j)^2}$$

Keterangan :

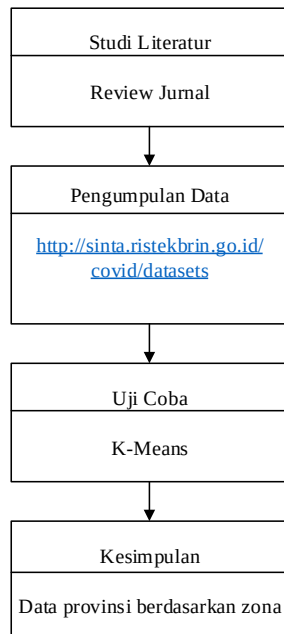
d = jarak  
j = banyaknya data  
c = centroid  
x = data

- Ulangi langkah 2 dan 3 sampai nilai dari titik *centroid* tidak lagi berubah.

Proses pengelompokan data ke dalam suatu *cluster* dapat dilakukan dengan cara menghitung jarak terdekat dari suatu data ke sebuah titik *centroid*. Perhitungan jarak Minkowski dapat digunakan untuk menghitung jarak antar 2 buah data. Rumus untuk menghitung jarak tersebut adalah (Maimon & Rokach, 2005):

## 3. Metodologi

Penelitian ini menggunakan jenis penelitian eksperimen, dengan tahapan penelitian seperti pada Gambar 2.



Gambar 2. Diagram Alir Penelitian

Adapun langkah-langkah yang dilakukan penulis untuk melakukan penelitian ini, yaitu :

#### 1. Studi Literatur

Pada tahap studi literatur, peneliti melakukan *review* jurnal yang berkaitan dengan pengelompokan menggunakan algoritma *K-Means* dengan kasus COVID-19. Jurnal yang menjadi acuan terdiri dari 3 (tiga) jurnal untuk mendapatkan keterbaruan dalam penelitian ini.

#### 2. Pengumpulan Data

Data yang digunakan dalam penelitian merupakan data yang diambil dari website

<http://sinta.ristekbrin.go.id/covid/datasets> yang berisi 34 provinsi yang ada di Indonesia. Dataset diambil pada tanggal 21 November 2020, dengan data seperti ditunjukkan pada Tabel 8.

Tabel 8  
Data Sebaran Per Provinsi

| No | Provinsi Asal   | Kasus   | Sembuh  | Kematian |
|----|-----------------|---------|---------|----------|
| 1  | Aceh            | 8,106   | 6,410   | 306      |
| 2  | Bali            | 13,012  | 11,922  | 406      |
| 3  | Banten          | 11,513  | 8,211   | 305      |
| 4  | Bangka Belitung | 873     | 754     | 11       |
| 5  | Bengkulu        | 1,538   | 1,179   | 66       |
| 6  | DI Yogyakarta   | 5,060   | 3,848   | 119      |
| 7  | DKI Jakarta     | 125,822 | 114,770 | 2,509    |

|    |                     |        |        |       |
|----|---------------------|--------|--------|-------|
| 8  | Jambi               | 1,620  | 1132   | 34    |
| 9  | Jawa Barat          | 47,692 | 38,318 | 865   |
| 10 | Jawa Tengah         | 46,903 | 34,818 | 2,110 |
| 11 | Jawa Timur          | 58,384 | 51,696 | 4,125 |
| 12 | Kalimantan Barat    | 2,231  | 1,620  | 22    |
| 13 | Kalimantan Timur    | 18,144 | 15,433 | 551   |
| 14 | Kalimantan Tengah   | 5,229  | 4,341  | 177   |
| 15 | Kalimantan Selatan  | 12,827 | 11,719 | 511   |
| 16 | Kalimantan Utara    | 1,088  | 865    | 13    |
| 17 | Kepulauan Riau      | 4,766  | 3,550  | 124   |
| 18 | Nusa Tenggara Barat | 4,511  | 3,590  | 234   |
| 19 | Sumatera Selatan    | 9,023  | 7,330  | 484   |
| 20 | Sumatera Barat      | 18,364 | 14,720 | 344   |
| 21 | Sulawesi Utara      | 6,399  | 5,076  | 231   |
| 22 | Sumatera Utara      | 14,831 | 12,219 | 591   |
| 23 | Sulawesi Tenggara   | 6,039  | 4,630  | 94    |
| 24 | Sulawesi Selatan    | 19,799 | 17,826 | 482   |
| 25 | Sulawesi Tengah     | 1,453  | 935    | 58    |
| 26 | Lampung             | 2,981  | 1,575  | 115   |
| 27 | Riau                | 18,471 | 16,014 | 409   |
| 28 | Maluku Utara        | 2,344  | 1,966  | 79    |
| 29 | Maluku              | 4,245  | 3,649  | 60    |
| 30 | Papua Barat         | 4,923  | 4,309  | 79    |
| 31 | Papua               | 9,759  | 5,007  | 137   |
| 32 | Sulawesi Barat      | 1,364  | 985    | 20    |
| 33 | Nusa Tenggara Timur | 937    | 630    | 14    |
| 34 | Gorontalo           | 3,057  | 2,908  | 89    |

#### 3. Uji Coba

Proses uji coba dilakukan dengan menggunakan *tool Rapid Miner*. Penelitian menggunakan algoritma *K-Means* untuk mengelompokkan provinsi kedalam 2 (dua) kriteria.

#### 4. Kesimpulan

Hasil uji coba menghasilkan 2 (dua) kelompok yang menunjukkan jumlah zona. Keanggotaan tiap kelompok merupakan provinsi yang memiliki kriteria yang mendekati dengan titik pusat yang merupakan nilai dari tiap zona.

#### 4. Hasil dan Pembahasan

Hasil penelitian menunjukkan bahwa algoritma *K-Means* menghasilkan 2 (dua) kelompok dengan keanggotaan *cluster* 0 sebanyak 6 data, *cluster* 1 sebanyak 28 data.

Pengelompokan data menggunakan algoritma *K-Means* dengan langkah sebagai berikut:

1. Pemilihan titik pusat/*centorid* secara acak. Pada kasus ini menggunakan data ke:

Tabel 9. Tabel Centoroid

| Data Ke | Provinsi Asal | Kasus   | Sembuh  | Kematian | Cluster |
|---------|---------------|---------|---------|----------|---------|
| 7       | DKI Jakarta   | 125,822 | 114,770 | 2,509    | c1      |
| 8       | Jambi         | 1,620   | 1132    | 34       | c2      |

2. Menghitung jarak data dengan *centroid*  
Jarak cluster

Misal data kota Aceh dengan kasus 8.106, sembuh 6.410, dan kematian 306, maka jarak data dengan centroid dapat dihitung sebagai berikut:

$$d_{x_1, c_1} = \sqrt{(8.106 - 125.822)^2 + (6.410 - 114.770)^2 + (306 - 2.509)^2} = 343.0827$$

$$d_{x_1, c_2} = \sqrt{(8.106 - 1.620)^2 + (6.410 - 1132)^2 + (306 - 34)^2} = 1158.006$$

Secara keseluruhan hasil perhitung data dengan *centroid* terlihat pada Tabel 9.

Tabel 9  
Hasil Perhitungan Jarak Ke-1

| No | Provinsi Asal       | Kasus   | Sembuh | Kematian | dc1      | dc2      | Cluster |
|----|---------------------|---------|--------|----------|----------|----------|---------|
| 1  | Aceh                | 8.106   | 6.41   | 306      | 343,0827 | 1158,006 | 1       |
| 2  | Bali                | 13,012  | 11,922 | 406      | 431,4033 | 1180,292 | 1       |
| 3  | Banten              | 11,513  | 8,211  | 305      | 340,4735 | 1156,045 | 1       |
| 4  | Bangka Belitung     | 873     | 754    | 11       | 983,3423 | 950,1137 | 2       |
| 5  | Bengkulu            | 1,538   | 1,179  | 66       | 179,9459 | 1131,274 | 1       |
| 6  | DI Yogyakarta       | 5,06    | 3,848  | 119      | 201,14   | 1131,355 | 1       |
| 7  | DKI Jakarta         | 125,822 | 114,77 | 2,509    | 0        | 1025,268 | 1       |
| 8  | Jambi               | 1,62    | 1132   | 34       | 1025,268 | 0        | 2       |
| 9  | Jawa Barat          | 47,692  | 38,318 | 865      | 869,3906 | 1374,345 | 1       |
| 10 | Jawa Tengah         | 46,903  | 34,818 | 2,11     | 112,3418 | 1098,579 | 1       |
| 11 | Jawa Timur          | 58,384  | 51,696 | 4,125    | 92,35164 | 1082,207 | 1       |
| 12 | Kalimantan          | 2,231   | 1,62   | 22       | 168,6937 | 1130,444 | 1       |
| 13 | Kalimantan Timur    | 18,144  | 15,433 | 551      | 567,7189 | 1230,562 | 1       |
| 14 | Kalimantan Tengah   | 5,229   | 4,341  | 177      | 239,1325 | 1136,696 | 1       |
| 15 | Kalimantan Selatan  | 12,827  | 11,719 | 511      | 530,9901 | 1217,655 | 1       |
| 16 | Kalimantan Utara    | 1,088   | 865    | 13       | 760,6009 | 267,8251 | 2       |
| 17 | Kepulauan Riau      | 4,766   | 3,55   | 124      | 204,4126 | 1132,038 | 1       |
| 18 | Nusa Tenggara Barat | 4,511   | 3,59   | 234      | 284,0166 | 1146,001 | 1       |
| 19 | Sumatera Selatan    | 9,023   | 7,33   | 484      | 506,9704 | 1211,378 | 1       |
| 20 | Sumatera Barat      | 18,364  | 14,72  | 344      | 371,7167 | 1159,61  | 1       |
| 21 | Sulawesi Utara      | 6,399   | 5,076  | 231      | 280,1834 | 1144,023 | 1       |
| 22 | Sumatera Utara      | 14,831  | 12,219 | 591      | 607,5832 | 1250,733 | 1       |
| 23 | Sulawesi Tenggara   | 6,039   | 4,63   | 94       | 186,6799 | 1128,974 | 1       |
| 24 | Sulawesi Selatan    | 19,799  | 17,826 | 482      | 500,5503 | 1201,007 | 1       |
| 25 | Sulawesi Tengah     | 1,453   | 935    | 58       | 831,4591 | 198,4566 | 2       |
| 26 | Lampung             | 2,981   | 1,575  | 115      | 201,3883 | 1133,324 | 1       |
| 27 | Riau                | 18,471  | 16,014 | 409      | 431,8703 | 1177,427 | 1       |
| 28 | Maluku Utara        | 2,344   | 1,966  | 79       | 183,9088 | 1130,93  | 1       |
| 29 | Maluku              | 4,245   | 3,649  | 60       | 174,4536 | 1128,654 | 1       |
| 30 | Papua Barat         | 4,923   | 4,309  | 79       | 180,7459 | 1128,593 | 1       |
| 31 | Papua               | 9,759   | 5,007  | 137      | 208,8214 | 1131,719 | 1       |
| 32 | Sulawesi Barat      | 1,364   | 985    | 20       | 879,2588 | 147,6654 | 2       |
| 33 | Nusa Tenggara Timur | 937     | 630    | 14       | 961,043  | 1061,763 | 1       |
| 34 | Gorontalo           | 3,057   | 2,908  | 89       | 187,2566 | 1130,432 | 1       |

Untuk menentukan keanggota *cluster*, dilihat dari jarak yang terdekat terhdap *cluster*.

### 3. Perbarui Centoroid

Jumlahkan seluruh anggota masing-masing *cluster* dan dibagi jumlah anggotanya.

Rumus :

$$\text{Cluster} = \frac{\text{Jumlah data anggota cluster}}{\text{banyak golongan cluster}} = \text{centeroid baru}$$

Menghitung *centroid* baru pada *cluster* 1

$$x = \frac{8,106 + 13,012 + 11,538 + 1,538 + 5,06 + 125,822 + 47,692 + 46,903 + 58,384 + 2,231 + 18,144 + 5,229 + 12,827 + 4,766 + 4,511 + 9,023 + 18,364 + 14,831 + 6,039 + 19,799 + 2,981 + 18,471 + 2,344 + 4,245 + 4,923 + 9,759 + 937 + 3,057}{29} = 49,06803$$

$$y = \frac{6,41 + 11,922 + 8,211 + 1,176 + 3,843 + 114,77 + 38,318 + 34,818 + 51,696 + 1,62 + 15,433 + 4,431 + 11,719 + 3,55 + 3,59 + 7,33 + 14,72 + 5,076 + 12,219 + 4,63 + 17,826 + 1,575 + 16,014 + 1,966 + 3,649 + 4,309 + 5,007 + 630 + 2,908}{29} = 35,81566$$

$$z = \frac{306 + 406 + 305 + 66 + 119 + 2,905 + 865 + 2,11 + 4,125 + 22 + 551 + 177 + 511 + 124 + 234 + 484 + 344 + 231 + 591 + 94 + 428 + 115 + 209 + 79 + 137 + 14 + 89}{29} = 238,0257$$

Menghitung *centeroid* baru pada *cluster* 2

$$x = \frac{873 + 1,62 + 1,088 + 1,453 + 1,364}{5} = 175,705$$

$$y = \frac{754 + 1132 + 865 + 935 + 985}{5} = 934,5$$

$$z = \frac{11 + 34 + 13 + 58 + 20}{5} = 27,2$$

Hasil perhitungan perbaikan *centroid* :  
Tabel 10 Perbaikan Centoroid Ke-1

| cluster 1 |          |          | cluster 2 |       |      |
|-----------|----------|----------|-----------|-------|------|
| X         | Y        | Z        | X         | Y     | Z    |
| 49,06803  | 35,81566 | 238,0257 | 175,705   | 934,2 | 27,2 |

4. Mengulangi menghitung jarak terdekat data centeroid, terlihat pada tabel 11.

Hitung kembali perbaikan centeroid pada Tabel 11.

Tabel 11

Hasil Perhitungan Jarak Ke-2

| No | Provinsi Asal | Kasus | Sembuh | Kematian | dc1 | dc2 | Cluster |
|----|---------------|-------|--------|----------|-----|-----|---------|
|----|---------------|-------|--------|----------|-----|-----|---------|

|    |                     |         |        |       |          |          |   |
|----|---------------------|---------|--------|-------|----------|----------|---|
| 1  | Aceh                | 8,106   | 6,41   | 306   | 343,0827 | 1158,006 | 1 |
| 2  | Bali                | 13,012  | 11,922 | 406   | 431,4033 | 1180,292 | 1 |
| 3  | Banten              | 11,513  | 8,211  | 305   | 340,4735 | 1156,045 | 1 |
| 4  | Bangka Belitung     | 873     | 754    | 11    | 983,3423 | 950,1137 | 2 |
| 5  | Bengkulu            | 1,538   | 1,179  | 66    | 179,9459 | 1131,274 | 1 |
| 6  | DI Yogyakarta       | 5,06    | 3,848  | 119   | 201,14   | 1131,355 | 1 |
| 7  | DKI Jakarta         | 125,822 | 114,77 | 2,509 | 0        | 1025,268 | 1 |
| 8  | Jambi               | 1,62    | 1132   | 34    | 1025,268 | 0        | 2 |
| 9  | Jawa Barat          | 47,692  | 38,318 | 865   | 869,3906 | 1374,345 | 1 |
| 10 | Jawa Tengah         | 46,903  | 34,818 | 2,11  | 112,3418 | 1098,579 | 1 |
| 11 | Jawa Timur          | 58,384  | 51,696 | 4,125 | 92,35164 | 1082,207 | 1 |
| 12 | Kalimantan Barat    | 2,231   | 1,62   | 22    | 168,6937 | 1130,444 | 1 |
| 13 | Kalimantan Timur    | 18,144  | 15,433 | 551   | 567,7189 | 1230,562 | 1 |
| 14 | Kalimantan Tengah   | 5,229   | 4,341  | 177   | 239,1325 | 1136,696 | 1 |
| 15 | Kalimantan Selatan  | 12,827  | 11,719 | 511   | 530,9901 | 1217,655 | 1 |
| 16 | Kalimantan Utara    | 1,088   | 865    | 13    | 760,6009 | 267,8251 | 2 |
| 17 | Kepulauan Riau      | 4,766   | 3,55   | 124   | 204,4126 | 1132,038 | 1 |
| 18 | Nusa Tenggara Barat | 4,511   | 3,59   | 234   | 284,0166 | 1146,001 | 1 |
| 19 | Sumatera Selatan    | 9,023   | 7,33   | 484   | 506,9704 | 1211,378 | 1 |
| 20 | Sumatera Barat      | 18,364  | 14,72  | 344   | 371,7167 | 1159,61  | 1 |
| 21 | Sulawesi Utara      | 6,399   | 5,076  | 231   | 280,1834 | 1144,023 | 1 |
| 22 | Sumatera Utara      | 14,831  | 12,219 | 591   | 607,5832 | 1250,733 | 1 |
| 23 | Sulawesi Tenggara   | 6,039   | 4,63   | 94    | 186,6799 | 1128,974 | 1 |
| 24 | Sulawesi Selatan    | 19,799  | 17,826 | 482   | 500,5503 | 1201,007 | 1 |
| 25 | Sulawesi Tengah     | 1,453   | 935    | 58    | 831,4591 | 198,4566 | 2 |
| 26 | Lampung             | 2,981   | 1,575  | 115   | 201,3883 | 1133,324 | 1 |
| 27 | Riau                | 18,471  | 16,014 | 409   | 431,8703 | 1177,427 | 1 |
| 28 | Maluku Utara        | 2,344   | 1,966  | 79    | 183,9088 | 1130,93  | 1 |
| 29 | Maluku              | 4,245   | 3,649  | 60    | 174,4536 | 1128,654 | 1 |
| 30 | Papua Barat         | 4,923   | 4,309  | 79    | 180,7459 | 1128,593 | 1 |
| 31 | Papua               | 9,759   | 5,007  | 137   | 208,8214 | 1131,719 | 1 |
| 32 | Sulawesi Barat      | 1,364   | 985    | 20    | 879,2588 | 147,6654 | 2 |
| 33 | Nusa Tenggara Timur | 937     | 630    | 14    | 961,043  | 1061,763 | 2 |
| 34 | Gorontalo           | 3,057   | 2,908  | 89    | 187,2566 | 1130,432 | 1 |

Menghitung *centroid* baru pada *cluster* 1

$$\begin{aligned}
 x &= \frac{8,106 + 13,012 + 11,538 + 1,538 + 5,06 + 125,822}{28} \\
 &+ \frac{47,692 + 46,903 + 58,384 + 2,231 + 18,144 + 5,229}{28} \\
 &+ \frac{12,827 + 4,766 + 4,511 + 9,023 + 18,364 + 6,399}{28} \\
 &+ \frac{14,831 + 6,039 + 19,799 + 2,981 + 18,471 + 2,344}{28} \\
 &+ \frac{4,245 + 4,923 + 9,759 + 3,057}{28} = 17,35618 \\
 y &= \frac{6,41 + 11,922 + 8,211 + 1,176 + 3,843 + 114,77}{28} \\
 &+ \frac{38,318 + 34,818 + 51,696 + 1,62 + 15,433 + 4,431}{28} \\
 &+ \frac{11,719 + 3,55 + 3,59 + 7,33 + 14,72 + 5,076 + 12,219}{28} \\
 &+ \frac{4,63 + 17,826 + 1,575 + 16,014 + 1,966 + 3,649}{28}
 \end{aligned}$$

$$\frac{+ 4,309 + 5,00 + 2,908}{28} = 14,59479$$

$$z = \frac{306 + 406 + 305 + 66 + 119 + 2,905 + 865 + 2,11}{28}$$

$$\frac{+ 4,125 + 22 + 551 + 177 + 511 + 124 + 234 + 484}{28}$$

$$\frac{+ 344 + 231 + 591 + 94 + 428 + 115 + 209 + 79}{28}$$

$$\frac{+ 60 + 79 + 137 + 89}{28} = 246,0266$$

Menghitung *centroid* baru pada *cluster* 2

$$x = \frac{873 + 1,62 + 1,088 + 1,453 + 1,364 + 937}{6} = 302,5875$$

$$y = \frac{754 + 1132 + 865 + 935 + 985 + 630}{6} = 883,5$$

$$z = \frac{11 + 34 + 13 + 58 + 20 + 14}{6} = 25$$

Kesimpulan dari perhitungan menunjukkan bahwa *centroid* tidak mengalami perubahan (sama dengan *centroid* sebelumnya) maka proses *clustering* selesai.

## 5. Simpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah menghasilkan perhitungan secara manual dan menggunakan *Rapid Miner* menunjukkan hasil yang sama. Penelitian menghasilkan 2 (dua) kelompok dengan keanggotaan *cluster* 0 menunjukkan wilayah yang paling sedikit terdampak covid-19 dengan jumlah 6 data, dan *cluster* 1 menunjukkan wilayah yang paling banyak terdampak covid-19 dengan jumlah 28 data. Dari penelitian ini maka dapat disimpulkan bahwa pemerintah sebaiknya melakukan penanganan covid-19 lebih terfokus pada *cluster* 1 yang jumlah terdampaknya paling banyak.

## Daftar Pustaka

- Dwitri, N., Tampubolon, J. A., Prayoga, S., R, F. I., & Hartama, D. (2020). Penerapan Algoritma *K-Means* Dalam Menentukan Tingkat Penyebaran Pandemi COVID-19 Di Indonesia. *Jurnal Teknologi Informasi*, 4(1), 101–105.



- Hammouda, K., & Karray, F. (2000). *A Comparative Study of Data Clustering Techniques*. *Course Project*, 2(1), 1–21.
- <https://covid19.go.id/peta-sebaran> 21 november 2020. (2020). Retrieved November 21, 2020, from <https://covid19.go.id/peta-sebaran> <https://pedulicovid19.kemenparekraf.go.id/pe-mbagian-zona-pada-pandemi-covid-19/>. (2020). Retrieved November 21, 2020, from <https://pedulicovid19.kemenparekraf.go.id/pe-mbagian-zona-pada-pandemi-covid-19/>
- <https://promkes.kemkes.go.id/informasi-tentang-virus-corona-novel-coronavirus>. (2020), (november), 2020.
- <https://www.kompas.com/sains/read/2020/05/11/130600623/diumumkan-awal-maret-ahli-virus-corona-masuk-indonesia-dari-januari>. (2020). Retrieved November 21, 2020, from <https://www.kompas.com/sains/read/2020/05/11/130600623/diumumkan-awal-maret-ahli-virus-corona-masuk-indonesia-dari-januari>
- Jain, A. K., & Dubes, R. C. (1988). *Algorithms for Clustering Data*. (B. Martine, Ed.). *New Jersey: Prentice Hall Advanced Reference Series*.
- Larose, D. T. (2005). *Discovering Knowledge In Data : an introduction to data mining*. In *Automotive Industries AI*. Canada: A John Wiley & Sons, Inc.
- Maimon, O., & Rokach, L. (2005). *Data mining and Knowledge Discovery*. (O. Maimon & L. Rokach, Eds.), *Springer Handbook of Geographic Information*. London: Springer New York Dordrecht Heidelberg.
- [https://doi.org/10.1007/978-3-540-72680-7\\_5](https://doi.org/10.1007/978-3-540-72680-7_5)
- Noviyanto. (2020). Penerapan Data Mining Dalam Mengelompokkan Jumlah Kematian. *Paradigma-Jurnal Informatika Dan Komputer*, 22(2).
- Nuryaman, Y., Asistiyasari, A., & Yudha, A. (2018). Komparasi Algoritma Kmean dan AHC Untuk Klasifikasi Curah Hujan di Indonesia. *IKRRAITH- INFORMATIKA*. Retrieved from <https://media.neliti.com/media/publications/226357-komparasi-algoritma-kmean-dan-ahc-untuk-597b4744.pdf>
- Yustanti, W., Rahmawati, N., & Yamasari, Y. (2020). Klastering Wilayah Kota/Kabupaten Berdasarkan Data Persebaran Covid-19 di Propinsi Jawa Timur dengan Metode K-Means. *Journal Information Engineering and Educational Technology*, 04(01), 1–9.