

**ANALISIS *K-MEANS* PADA PENGELOMPOKAN
KABUPATEN-KOTA PROVINSI JAWA TIMUR
BERDASARKAN KASUS KESEMBUHAN DAN KASUS
KEMATIAN COVID-19**

SKRIPSI

**OLEH
YENI DWI MAYANGSARI
NIM. 17610008**



**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2022**

**ANALISIS *K-MEANS* PADA PENGELOMPOKAN
KABUPATEN-KOTA PROVINSI JAWA TIMUR
BERDASARKAN KASUS KESEMBUHAN DAN KASUS
KEMATIAN COVID-19**

SKRIPSI

**Diajukan Kepada
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh
Yeni Dwi Mayangsari
NIM. 17610008**

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2022**

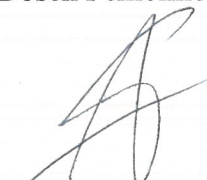
**ANALISIS K-MEANS PADA PENGELOMPOKAN
KABUPATEN-KOTA PROVINSI JAWA TIMUR
BERDASARKAN KASUS KESEMBUHAN DAN KASUS
KEMATIAN COVID-19**

SKRIPSI

**Oleh
Yeni Dwi Mayangsari
NIM. 17610008**

Telah Diperiksa dan Disetujui untuk Diuji
Malang, 14 Desember 2022

Dosen Pembimbing I



Angga Dwi Mulyanto, M.Si
NIP. 198908132019031012

Dosen Pembimbing II


Erna Herawati, M.Pd
NIP. 19760723201802012222



Mengetahui,
Ketua Program Studi Matematika


Dr. Elly Susanti, M.Sc
NIP. 197411292000122005

**ANALISIS *K-MEANS* PADA PENGELOMPOKAN
KABUPATEN-KOTA PROVINSI JAWA TIMUR
BERDASARKAN KASUS KESEMBUHAN DAN KASUS
KEMATIAN COVID-19**

SKRIPSI

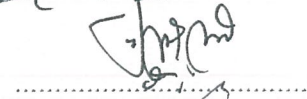
Oleh
Yeni Dwi Mayangsari
NIM. 17610008

Telah Dipertahankan di Depan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)
Malang, 21 Desember 2022

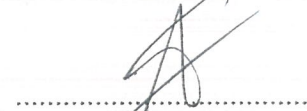
Ketua Penguji : Abdul Aziz, M.Si



Anggota Penguji 1 : Ria Dhe Layla N.K, M.Si



Anggota Penguji 2 : Angga Dwi Mulyanto, M.Si



Anggota Penguji 3 : Erna Herawati, M.Pd



Mengetahui,
Ketua Program Studi Matematika



Dr. Ely Susanti, M.Sc
NIP. 197411292000122005

PERNYATAAN KEASLIAN TULISAN

Saya yang bertandatangan dibawah ini:

Nama : Yeni Dwi Mayangsari

NIM : 17610008

Jurusan : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Analisis *K-means* Pada Pengelompokan Kabupaten-Kota Provinsi Jawa Timur Berdasarkan Kasus Kesembuhan Dan Kasus Kematian COVID-19

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan dan pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar rujukan. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 21 Desember 2022

Yang membuat pernyataan,




Yeni Dwi Mayangsari
NIM.17610008

MOTTO

فَإِذَا فَرَغْتَ فَانصَبْ

“Maka apabila kamu telah selesai (dari suatu urusan), kerjakanlah dengan
sungguh-sungguh (urusan) yang lain”

(QS. Al-Insyirah ayat 7)

PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dengan ucapan *Alhamdulillah* sebagai ungkapan rasa syukur kepada Allah SWT, skripsi penulis persembahkan teruntuk:

Bapak Yahman dan Ibu Djumilah, serta kakak Yuli Safitri yang menjadi motivasi utama penulis dalam menyelesaikan skripsi ini, serta sanak saudara, kerabat serta teman-teman yang senantiasa dengan ikhlas mendoakan, memberikan motivasi, nasehat, dukungan serta kasih sayang yang tak terhingga.

KATA PENGANTAR

Assalamualaikum Warrahmatullahi Wabarakatuh.

Puji Syukur penulis haturkan kehadiran Allah SWT yang senantiasa memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Analisis *K-means* Pada Pengelompokan Kabupaten-Kota Provinsi Jawa Timur Berdasarkan Kasus Kesembuhan Dan Kasus Kematian COVID-19” sebagai salah satu syarat untuk memperoleh gelar sarjana dalam bidang Matematika di Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Shalawat dan salam semoga tercurahkan kepada Nabi Muhammad SAW yang telah membimbing manusia dari zaman jahiliah menuju zaman islamiah.

Dalam proses penyusunan skripsi ini, tentunya penulis banyak menerima bantuan, bimbingan, masukan, dan arahan dari berbagai pihak. Oleh karena itu penulis sampaikan ucapan terimakasih yang sebesar-besarnya melalui halaman ini kepada:

1. Prof. Dr. H. M. Zainuddin, M.A, selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Elly Susanti, M.Sc, selaku ketua Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Angga Dwi Mulyanto, M.Si, selaku dosen pembimbing I yang telah memberikan banyak ilmu, arahan, nasehat serta motivasi kepada penulis.
5. Erna Herawati, M.Pd, selaku dosen pembimbing II yang telah memberikan banyak ilmu, arahan, nasehat serta arahan kepada penulis.
6. Seluruh Dosen dan civitas akademik Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim .
7. Orang tua dan seluruh keluarga yang selalu mengirimkan doa terbaik kepada penulis.
8. Seluruh teman-teman di Program Studi Matematika angkatan 2017.

9. Seluruh pihak yang tidak dapat disebutkan satu-persatu, yang telah ikut serta membantu menyelesaikan penyusunan skripsi, baik dukungan moril maupun materil.

Penulis sadar tidak bisa memberikan apapun selain ucapan terima kasih dan doa semoga Allah membalas kebaikan semua pihak dengan balasan yang sebaik-baiknya. Penulis berharap semoga skripsi ini dapat bermanfaat, baik bagi penulis maupun pembaca.

Wassalamualaikum Warahmatullahi Wabarakatuh.

Malang, 21 Desember 2022

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xiv
ABSTRAK	xv
ABSTRACT	xvi
مستخلص البحث	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
1.5 Batasan Masalah.....	6
1.6 Definisi Istilah	6
BAB II KAJIAN PUSTAKA	8
2.1 Analisis <i>Cluster</i>	8
2.2 Deteksi <i>Outlier</i>	9
2.3 Metode Perhitungan Jarak.....	10
2.4 Normalisasi Data	11
2.5 Uji Multikolinieritas	11
2.6 <i>K-Means</i>	14
2.7 Metode <i>Elbow</i>	16
2.8 COVID-19	18
2.9 Analisis <i>Cluster</i> dalam Islam	20
2.10 Analisis <i>Cluster</i> COVID-19 di Provinsi Jawa Timur.....	23
BAB III METODOLOGI PENELITIAN	24
3.1 Jenis Penelitian	24
3.2 Data dan Sumber Data.....	24
3.3 Lokasi Penelitian	24
3.4 Teknik Pengumpulan Data	25
3.5 Instrumen Penelitian.....	25
3.6 Teknik Analisis Data	26
BAB IV HASIL DAN PEMBAHASAN	28
4.1 Statiska Deskriptif	28
4.2 Deteksi <i>Outlier</i>	31
4.3 Uji Asumsi Analisis <i>Cluster</i>	32
4.4 Normalisasi Data	35

4.5	Penentuan Jumlah Klaster	35
4.6	Pengelompokan dan Interpretasinya	37
4.7	Ciri Kelompok.....	42
4.8	Kajian Keislaman dengan Hasil Penelitian	45
BAB V PENUTUP.....		47
5.1	Kesimpulan.....	47
5.2	Saran.....	48
DAFTAR PUSTAKA		
LAMPIRAN		

DAFTAR TABEL

Tabel 4.1 Statistika Deskriptif	28
Tabel 4.2 Perhitungan SST dan SSR.....	33
Tabel 4.3 Tabel Nilai SSE.....	36
Tabel 4.4 <i>Centroid</i> Awal.....	37
Tabel 4.5 <i>Centroid</i> Baru.....	39
Tabel 4.6 Hasil Pengklasteran <i>K-Means</i>	40
Tabel 4.7 Hasil Pengelompokan <i>K-Means Cluster</i>	41

DAFTAR GAMBAR

Gambar 2.1 Contoh Grafik Metode <i>Elbow</i>	17
Gambar 4.1 Grafik Penyebaran Kasus Sembuh COVID-19 di Provinsi Jawa Timur.....	29
Gambar 4.2 Grafik Penyebaran Kasus Meninggal Akibat COVID-19 di Provinsi Jawa Timur.....	30
Gambar 4.3 <i>Boxplot</i> X1.....	31
Gambar 4.4 <i>Boxplot</i> X2.....	31
Gambar 4.5 Grafik Metode <i>Elbow</i>	36
Gambar 4.6 <i>Scatterplot</i> Meninggal Vs Sembuh	43

DAFTAR LAMPIRAN

- Lampiran 1. Data Persebaran COVID-19 per 31 Juli 2021
- Lampiran 2. Matriks H
- Lampiran 3. Matriks J
- Lampiran 4. Matriks I
- Lampiran 5. Data setelah dinormalisasi
- Lampiran 6. Hasil Perhitungan *Euclidian Distance* pada iterasi-1
- Lampiran 7. Hasil Perhitungan *Euclidian Distance* pada iterasi-2
- Lampiran 8. Hasil Perhitungan *Euclidian Distance* pada iterasi-3
- Lampiran 9. Hasil Perhitungan *Euclidian Distance* pada iterasi-4
- Lampiran 10. Hasil Perhitungan *Euclidian Distance* pada iterasi-5
- Lampiran 11. Hasil Perhitungan *Euclidian Distance* pada iterasi-6
- Lampiran 12. Hasil Perhitungan *Euclidian Distance* pada iterasi-7
- Lampiran 13. Hasil Perhitungan *Euclidian Distance* pada iterasi-8
- Lampiran 14. *Script Python*

ABSTRAK

Mayangsari, Yeni Dwi. 2022. Analisis *K-means* pada Pengelompokan Kabupaten-Kota Provinsi Jawa Timur Berdasarkan Kasus Kesembuhan dan Kasus Kematian COVID-19. Skripsi. Jurusan Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Angga Dwi Mulyanto, M.Si. (II) Erna Herawati, M.Pd.

Kata kunci: analisis *cluster*, *k-means*, *elbow*, jarak *euclidean*, COVID-19 .

K-means Cluster merupakan satu metode non-hierarki yang mempartisi data yang ada menjadi beberapa *cluster*. Konsep penelitian ini menentukan banyak kelompok dengan menggunakan metode *elbow*. Hasil penelitian ini adalah mengelompokkan penyebaran kasus kesembuhan dan kematian akibat COVID-19 kabupaten-kota di Provinsi Jawa Timur dan menentukan ciri dari setiap kelompok yang terbentuk. Hasil dari penelitian ini diperoleh $K = 4$ sebagai nilai *optimal* yang diperoleh dari nilai SSE pada grafik *elbow*. Pengelompokan iterasi *k-means* dilakukan sebanyak 8 kali dengan hasil pengelompokan yaitu *cluster* 0 dengan kasus sangat tinggi sebanyak 1 kota, *cluster* 1 dengan kasus tinggi sebanyak 9 kabupaten-kota, *cluster* 2 dengan kasus rendah sebanyak 13 kabupaten-kota, dan *cluster* 3 dengan kasus paling rendah sebanyak 15 kabupaten-kota.

ABSTRACT

Mayangsari, Yeni Dwi. 2022. K-means Analysis on the Grouping of Regencies and Cities of East Java Province Based on COVID-19 Recovery and Death Cases. Thesis. Department of Mathematics, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
Supervisor: (I) Angga Dwi Mulyanto, M.Si. (II) Erna Herawati, M.Pd.

Keywords: Cluster analysis, k-means, elbow, Euclidean Distance, COVID-19.

K-means Cluster is a non-hierarchical method that partitions existing data into multiple clusters. The concept of this research determined many groups using the elbow method. The results of this study are to group the spread of cases of recovery and death due to COVID-19 in East Java Province and determine the characteristics of each group formed. The results of this study obtained the amount of $K = 4$ as the optimal value obtained from the SSE value on the elbow graph. The grouping of k-means iterations was carried out 8 times with the results of grouping cluster 1 with very high cases as many as 1 city, cluster 2 with high cases as many as 9 districts-cities, cluster 3 with low cases with 13 districts, and cluster 4 with the lowest cases as many as 15 districts-cities.

مستخلص البحث

مايانع ساري، يني دووي. ٢٠٢٢. التحليل *k-means* في تصنيف منطقة-مدينة في جاوى الشرقية بناءً على حالة الشفاء و الوفاة من كوفيد - 19. البحث الجامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية بمالانج. المشرف: (١) أنغا دوي موليانو ، الماجستير. (٢) إيرنا هيراواتي ، الماجستير.

الكلمات الرئيسية: تطبيقات التصنيف، *k-means*، *elbow*، المسافة الإقليدية، كوفيد - 19 .

الخوارزمية التصنيفية (*k-means clustering*) هي طريقة ليس من سلسلة المراتب التي تقسم عدد من البيانات الموجودة الى عدد من الأقسام. يحدد مفهوم هذا البحث عدد المجموعات باستخدام طريقة *elbow*. النتائج هذا البحث هي تصنيف انتشار حالة الشفاء والوفاة بسبب كوفيد - 19 في منطقة-مدينة في جاوى الشرقية و تحديد خصائص كل مجموعة تم تشكيلها. يتم الحصول من هذا البحث على قيمة $K = 4$ كالقيمة المثلى التي تم الحصول عليها من قيمة *SSE* على مخطط *elbow*. تم إجراء تجميع تكرارات *k-means* ثمانية مرات مع نتائج التجميع وهي المجموعة 0 مع مدينة واحدة ذات حالات عالية جدًا ، والمجموعة 1 مع تسعة مناطق ذات حالة عالية، والمجموعة 2 مع ثلاثة عشر مناطق ذات حالة منخفضة ، والمجموعة 3 مع خمسة عشر مناطق-مدن ذات اقل عدد حالات.

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis *cluster* adalah metode analisa statistik multivariat yang bertujuan mengelompokkan objek yang karakteristiknya sama, sehingga objek didalam kelompok serupa mempunyai sifat yang agak homogen jika dibandingkan dengan kelompok lain. Seharusnya, ada tingkat keseragaman (kesamaan) yang tinggi diantara anggota dari satu kelompok (*within-cluster*) dan tingkat heterogenitas yang tinggi antara satu kelompok dengan kelompok lainnya (*between-cluster*). Sistem pengelompokan, analisis *cluster* pada umumnya dipisahkan menjadi metode *hierarki* dan *non hierarki* .

Pengelompokan menggunakan metode hierarki digunakan ketika tidak terdapat informasi tentang jumlah *cluster*, sedangkan pengelompokan dengan metode non hierarki digunakan ketika jumlah *cluster* diketahui atau ditentukan sebagai bagian dari prosedur pengelompokan (Johnson & Winchern, 2007). Pada teknik nonhierarki dimulai dengan menentukan jumlah *cluster* yang *optimal* terlebih dahulu, kemudian sistem pengelompokan selesai tanpa mengikuti proses hirarkinya. Teknik ini dikenal sebagai *k-means cluster*.

Penelitian sebelumnya yang membahas tentang pengelompokan *k-means* yakni seperti pada penelitian (Rahman, Wiranto, & Anggrainingsih, 2017) yakni “*Coal Trade Data Clustering Using K-means (Case Study PT.Global Bangkit Umum)*”. Penelitian tersebut bertujuan membuat visualisasi *cluster optimal* dengan menggunakan metode *elbow* pada data penjualan batubara PT Global Bangkit Utama yang menghasilkan *cluster optimal* sebanyak 8 *cluster*. Selain itu, pada penelitian (Zurraedah, 2018) yakni “Evaluasi Klasifikasi Kategori

Pembayaran SPP di Universitas Islam Negeri Alauddin Makassar (UINAM) Tahun Akademik 2017/2018 Menggunakan *K-Means Clustering*”. Pada penelitian tersebut dilakukan pengelompokan *k-means* menggunakan *euclidean distance* pada data Mahasiswa Universitas Islam Negeri Alauddin Makassar (UINAM) mengenai pembayaran SPP kemudian didapatkan hasil pengelompokan UKT yang berbeda yang ditetapkan oleh Perguruan Tinggi Negeri (PTN) UINAM dengan hasil evaluasi *k-means clustering*

Pengelompokan dalam teknik *k-means* memiliki ketelitian yang cukup tinggi terhadap ukuran objek, oleh karena itu algoritma ini relatif lebih terukur dan efisien untuk menangani objek yang jumlahnya banyak. Cara utama dalam proses penghitungan *k-means* adalah dengan mengenali jumlah kelompok *k*, dimana hasil *cluster* berketergantungan terhadap jumlah *cluster* awalnya. Maka apabila dalam menentukan jumlah *cluster* tidak baik, akibatnya hasil *cluster* juga tidak akan sesuai dengan yang diharapkan yakni tidak menghasilkan informasi yang dibutuhkan oleh pengguna (Ernawati, Merliana, & Santoso, 2015). Adapun untuk mengatasi hal tersebut peneliti menggunakan metode *elbow* guna mencari nilai *k optimal* pada pengelompokan *k-means*.

Metode *elbow* adalah teknik yang diperuntukkan menghasilkan informasi data dalam menentukan jumlah kelompok terbaik dengan memeriksa tingkat hasil antara perbandingan jumlah *cluster* pada titik yang membentuk siku. Hasil dari *k cluster* yang *optimal* tersebut kemudian dijadikan alasan sistem pengumpulan menggunakan *k-means* (Rahman, Wiranto, & Anggrainingsih, 2017).

COVID-19 atau yang disebut juga virus corona merupakan sekumpulan virus yang menyerang sistem pernapasan ringan hingga berat, seperti flu hingga

infeksi paru-paru, juga penyakit serius seperti MERS dan SARS (Yunus & Rezki, 2020). Pertama kali virus tersebut ditemukan di kota Wuhan, Provinsi Hubei, Tiongkok pada akhir Desember 2019. Dalam Al-Qur'an telah dijelaskan mengenai virus, salah satunya pada QS.Al-Baqarah ayat 26 yakni:

إِنَّ اللَّهَ لَا يَسْتَحْيِي أَنْ يَضْرِبَ مَثَلًا مَّا بَعُوضَةً فَمَا فَوْقَهَا فَأَمَّا الَّذِينَ آمَنُوا فَيَعْلَمُونَ أَنَّهُ الْحَقُّ مِنْ رَبِّهِمْ وَأَمَّا الَّذِينَ كَفَرُوا فَيَقُولُونَ مَاذَا أَرَادَ اللَّهُ بِهَذَا مَثَلًا يُضِلُّ بِهِ كَثِيرًا وَيَهْدِي بِهِ كَثِيرًا وَمَا يُضِلُّ بِهِ إِلَّا الْفَاسِقِينَ (٢٦)

Artinya:

“Sesungguhnya Allah tidak segan membuat perumpamaan seekor nyamuk atau yang lebih kecil dari itu. Adapun orang-orang yang beriman, maka mereka tahu bahwa itu kebenaran dari Tuhan, tetapi mereka yang kafir berkata: "Apakah maksud Allah dengan perumpamaan ini?". Dengan (perumpamaan) itu banyak orang yang disesatkanNya, dan dengan perumpamaan itu banyak (pula) orang yang diberi-Nya petunjuk. Tetapi tidak ada yang Dia sesatkan selain orang-orang yang fasik”.

Menurut Tafsir Al Misbah (Shihab, 2002) ayat tersebut menjelaskan bahwasannya Allah tidak segan membuat perumpamaannya untuk kebenaran dengan seekor nyamuk ataupun sesuatu kecil yang melebihinya, bahkan kecil sekali hingga tak kasat mata namun mampu menggemparkan manusia dengan ketakutan dan kepanikan diberbagai belahan dunia, seperti COVID-19 yang hanya memiliki ukuran 125 nanometer atau 0.125 mikrometer. Perumpamaan makhluk yang sangat kecil tersebut menunjukkan tentang keunikan dan keajaiban ciptaanNya yang tidak pantas disepelakan oleh manusia. Seseorang yang beriman akan terus memohon kepada Allah untuk keselamatan dan menggunakan akal nya untuk mencari tahu, mencari solusi dari kasus penyebaran COVID-19 ini, sebab setiap penyakit belum tentu mendatangkan kematian. Sedangkan orang yang fasik yakni orang yang telah berkurang imannya atau enggan untuk mendekatkan diri kepadaNya justru enggan untuk ikhtiar dan melalui masa pandemi ini dengan

angkuk, seperti enggan mematuhi intruksi dari pihak yang berwenang dan tidak mematuhi protokol kesehatan, sehingga menyebabkan keadaan masa pandemi semakin kacau dan tidak mampu mengambil pelajaran dari perumpamaan-perumpamaan tersebut.

Penyebaran virus COVID-19 relatif cepat dan meluas karena transmisi virus COVID-19 dapat melalui kontak manusia dengan manusia. Pada tanggal 12 Maret 2020 lalu, *World Health Organization* mengabarkan COVID-19 sebagai pandemi sebab dampaknya menyebar yang meluas pada saat itu (WHO, 2020). Pada laman Jatim Tanggap COVID-19 pada 10 Maret 2020, diinformasikan Provinsi Jawa Timur mencapai 133.289 kasus terkonfirmasi COVID-19, dengan 121.321 kasus sembuh, 2.584 dirawat dan meninggal 9.384 jiwa (Pemprov Jatim, 2021). Pada tahun 2016, Penduduk Provinsi Jawa Timur jumlahnya adalah 39.0750152 jiwa dengan jumlah penduduk laki-laki 19.288.006 jiwa dan penduduk perempuan 19.787.146 jiwa. Jumlah penduduk yang cukup padat, sehingga menjadikan Provinsi Jawa Timur termasuk salah satu provinsi di Indonesia yang mempunyai tingkat infeksi dan tingkat kematian yang relatif tinggi (Soetcipto, 2020) .

Dalam mengurangi penularan COVID-19 agar tidak semakin meningkat upaya antisipasi sangat dibutuhkan. Upaya ini dapat dilakukan salah satunya dengan cara mengetahui karakteristik kasus COVID-19 di kabupaten-kota Provinsi Jawa Timur kemudian mengelompokkan wilayah kabupaten-kota yang memiliki karakteristik yang sama. Sehingga dapat diketahui karakteristik kelompok pada klaster yang terbentuk dengan *k-means cluster* untuk setiap wilayah masing-masing klaster.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana hasil pengelompokan analisis *k-means* pada kabupaten-kota Provinsi Jawa Timur menurut kasus kesembuhan dan kasus kematian COVID-19?
2. Bagaimana ciri dari setiap kelompok yang terbentuk?

1.3 Tujuan Penelitian

Tujuan pada penelitian ini adalah sebagai berikut:

1. Mengetahui hasil pengelompokan analisis *k-means* pada kabupaten-kota Provinsi Jawa Timur menurut kasus kesembuhan dan kasus kematian COVID-19.
2. Mengetahui ciri dari setiap kelompok yang terbentuk.

1.4 Manfaat Penelitian

Adapun manfaat yang dapat diberikan dalam penelitian adalah sebagai berikut:

1. Bagi peneliti

Dapat menambah wawasan mengenai aplikasi analisis *cluster* pada pengelompokan kasus COVID-19 di Provinsi Jawa Timur.

2. Bagi pembaca

Dapat memberikan wawasan pada bidang matematika khususnya mengenai analisis *cluster*.

3. Bagi Instansi

Dapat digunakan sebagai tambahan kepustakaan dalam pengembangan ilmu pengetahuan Matematika, khususnya bagi mahasiswa Jurusan Matematika.

1.5 Batasan Masalah

Batasan masalah pada penelitian ini yakni data yang digunakan merupakan data pasien sembuh dan meninggal akibat COVID-19 Provinsi Jawa Timur periode 31 Juli 2021 yang diambil dari laman Jatim Tanggap COVID-19 (Pemprov Jatim, 2021).

1.6 Definisi Istilah

Berikut beberapa definisi istilah pada penelitian ini agar lebih mudah dipahami oleh pembaca:

Analisis cluster : Salah satu teknik analisis yang bertujuan untuk mengelompokkan objek-objek kedalam suatu yang memiliki kesamaan karakteristik antar objek satu dengan yang lainnya.

Euclidean Distance : Perhitungan jarak antara dua titik dengan menggunakan *euclidean space*

Centroid : Titik tengah yang merupakan nilai rata-rata dalam suatu *cluster*.

Iterasi : Pengulangan fungsi atau program.

Koefisien korelasi parsial : Bilangan yang menunjukkan kekuatan hubungan antar dua variabel.

Tranpose : Suatu matriks yang diperoleh dari hasil pertukaran antara elemen baris dan kolomnya.

BAB II KAJIAN PUSTAKA

2.1 Analisis *Cluster*

Analisis *cluster* adalah sekelompok teknik multivariat yang bertujuan mengelompokkan objek berdasarkan karakteristik yang dimiliki. Analisis *cluster* mengelompokkan objek data (seperti data responden, data produk, dan sebagainya) menjadi satu kelompok yang memiliki kemiripan homogenitas (dalam *cluster*) dan heterogenitas (antar *cluster*) yang tinggi. Sehingga ketika pengelompokan telah berhasil dilakukan, maka objek-objek didalam *cluster* akan berdekatan dan *cluster* yang berbeda akan terpisah jika ditinjau dari plot yang dihasilkan (Hair, dkk., 2009).

Sedangkan menurut (Han & Kamber, 2001), analisis *cluster* adalah teknik data mining yang mengelompokkan sekumpulan objek (dataset) menjadi beberapa *cluster* saja sesuai dengan kesamaan karakteristik dan atribut dari data objek tersebut, sehingga data objek berada didalam *cluster* yang sama memiliki keserupaan satu sama lain tetapi tidak serupa dengan data objek yang berada dalam *cluster* yang lain. Hasil dari analisis *cluster* dipengaruhi oleh pengelompokan objeknya, variabel yang diamati, ukuran kemiripan atau perbedaan yang digunakan, skala ukuran yang digunakan dan teknik pengelompokan yang digunakan. Adapun ciri kelompok yang baik yakni (Prakasa & Kurnianingtyas, 2022):

1. *Homogenitas internal (within cluster)*, yaitu kesamaan antar individu (anggota) dalam satu kelompok. Misalnya, sekelompok *customer* ibu rumah tangga yang peduli dengan lingkungannya. Pasti yang termasuk dalam kelompok ini

adalah orang-orang yang memiliki kecintaan dan mengutamakan terhadap kebersihan dan kenyamanan lingkungan rumah, kemudian *customer* lainnya semisal *customer* yang lebih memilih untuk membeli harga rumah yang murah, pastinya tidak masuk kedalam kelompok *customer* yang mengutamakan kenyamanan lingkungan.

2. Heterogenitas luar (*between cluster*), yaitu perbedaan antara satu kelompok dengan kelompok lainnya. Dalam permasalahan *customer customer* ibu rumahtangga yang peduli dengan lingkungannya, maka anggota dari *cluster* ibu rumahtangga yang mengutamakan lingkungan memiliki prespektif yang berbeda dibandingkan dengan anggota *customer* yang lebih mengutamakan rumah dengan harga yang murah.

Teknik dalam melakukan sistem pengelompokan, diantaranya: menggunakan metode *hierarki* dan *non hierarki*. Metode *hierarichal* mengelompokkan data kedalam hierarki atau pohon *cluster*, sedangkan *metode nonhierarki (partisi)* mengelompokkan objek data kedalam beberapa *cluster*. Metode yang sering digunakan salah satunya adalah metode *k-means*.

2.2 Deteksi *Outlier*

Sebelum melakukan analisis *cluster*, perlu dilakukan deteksi *outlier* (pencilan). Deteksi *outlier* dilakukan karena analisis *cluster* sensitif terhadap variabel yang tidak relevan. *Outlier* dapat berupa pengamatan yang benar-benar menyimpang yang tidak mewakili populasi umum dan tidak merepresentatif kelompok yang sebenarnya dalam populasi. Dalam kasus pertama, *outlier* melakukan penyimpangan dari suatu data yang benar-benar terjadi. Dalam kasus

kedua, *outlier* dapat dihilangkan sehingga *cluster* yang dihasilkan lebih akurat. Akan tetapi pada kasus ketiga *outlier* tetap dimasukkan dalam analisis *cluster*, bahkan jika *outlier* kurang terwakili dalam sampel, karena *outlier* mewakili kelompok yang valid dan relevan. Oleh karena itu deteksi *outlier* pada tahap awal perlu dilakukan (Hair, dkk., 2009).

2.3 Metode Perhitungan Jarak

Untuk membentuk kelompok perlu diketahui tingkat kemiripan dari objek satu dengan objek lainnya. Tingkat kemiripan yang paling umum digunakan adalah ukuran jarak. Ukuran jarak ini mewakili kemiripan atau kedekatan objek satu dengan lainnya. Semakin kecil jarak yang dimiliki antara objek satu dengan objek lain maka dapat dikatakan kemiripan dari objek-objek tersebut sangat tinggi, sehingga berkumpul dalam 1 *cluster*. Dalam analisis *cluster* terdapat beberapa cara untuk melakukan perhitungan jarak salah satunya adalah *euclidean distance* (Hair, dkk., 2009).

Jarak *euclidean* merupakan jarak rata-rata dari 2 objek. Perhitungan jarak *euclidean* sama dengan perhitungan pythagoras, dimana untuk mendapatkan ukuran sisi miring maka perlu dicari dengan mencari akar dari jumlah kedua sisi siku yang telah dikuadratkan. Berikut rumus *euclidean distance* (Johnson & Winchern, 2007):

$$d(X.Y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (2.1)$$

dimana:

$d(X.Y)$ = Jarak antara objek X dan objek Y

n = Jumlah variabel *cluster*

x_n, y_n = Data ke-n dari objek X dan Y

2.4 Normalisasi Data

Standarisasi data berfungsi untuk menghilangkan bias yang disebabkan oleh perbedaan satuan variabel yang digunakan pada proses pengelompokan. Standarisasi data perlu dilakukan untuk menghindari pengaruh antara variabel satu dengan variabel yang lainnya (Hair, dkk., 2009). *Min-max Normalization* merupakan salah satu metode standarisasi data yang dilakukakukan agar data berada pada interval yang lebih kecil seperti $[-1,1]$ atau $[0,1]$, sehingga tidak ada data yang terlalu besar maupun terlalu kecil untuk memudahkan analisis. Metode ini merubah nilai data menjadi nilai dalam skala minimal dan maksimal yang diinginkan. Berikut adalah rumus untuk melakukan *min-max normalization* (Berry,2016):

$$x_{baru} = \frac{x_{lama} - x_{min}}{x_{max} - x_{min}} \quad (2. 2)$$

Keterangan:

x_{baru} = Nilai data hasil normalisasi

x_{lama} = Nilai data awal sebelum dinormalisasi

x_{min} = Nilai minimum pada data awal

x_{max} = Nilai maksimum pada data awal

2.5 Uji Multikolinieritas

Analisis *cluster* bukan merupakan teknik inferensi statistik dimana parameter dari sampel dapat dikatakan mewakili populasi. Analisis *cluster*

merupakan metode untuk mengukur karakteristik struktural dari beberapa pengamatan. Pada penelitian lain uji normalitas, uji multikolinieritas dan homoskedasitas perlu dilakukan. Namun berbeda dengan analisis *cluster*, uji yang dapat dilakukan sebelum melakukan pengelompokan ada dua macam. Uji kecukupan sampel dan uji multikolinieritas (Hair, dkk., 2009)., namun pada penelitian ini hanya uji multikolinieritas yang dilakukan sebab data yang digunakan merupakan data populasi. Multikolinieritas adalah adanya hubungan pada suatu obyek atau korelasi pada suatu obyek. Untuk melakukan proses pengelompokan sebaiknya tidak terdapat korelasi antar variabel sehingga dianjurkan untuk mengeluarkan variabel-variabel yang memiliki korelasi cukup besar. Multikolinieritas dapat diketahui dengan mencari nilai VIF (*Variance Inflation Factor*). Berikut adalah langkah-langkah mencari nilai VIF (Relia Soft Corporation, 2018):

1. Membuat matriks identitas dengan ukuran sebanyak $n \times n$, yakni:

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} \quad (2.3)$$

dan juga matriks yang seluruh elemennya bernilai 1 dengan ukuran sebanyak $n \times n$

$$J = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \end{bmatrix} \quad (2.4)$$

dimana n adalah banyaknya data. Selanjutnya,

$$H = X(X'X)^{-1}X' \quad (2.5)$$

Keterangan:

X = Matriks X

X' = Tranpose matriks X

$(X'X)^{-1}$ = Invers dari $(X'X)$

2. Mencari jumlah kuadrat regresi dan jumlah kuadrat total, dimana:

$$SS_R = y' \left[H - \left(\frac{1}{n} \right) J \right] y \quad (2.6)$$

Keterangan:

SS_R = Jumlah kuadrat regresi

y' = Tranpose y

n = Banyaknya data

y = Variabel dependen

J = Matriks 1 dengan ordo nxn

$$SS_T = y' \left[I - \left(\frac{1}{n} \right) J \right] y \quad (2.7)$$

Keterangan:

SS_T = Jumlah kuadrat total

y' = Tranpose y

n = Banyaknya data

y = Variabel dependen

I = Matriks Identitas dengan ordo nxn

J = Matriks 1 dengan ordo nxn

3. Menghitung nilai koefisien determinasi

$$R^2 = \frac{SS_R}{SS_T} \quad (2.8)$$

Keterangan:

R^2 = Koefisien determinasi

SS_R = Jumlah kuadrat regresi

SS_T = Jumlah kuadrat total

4. Menentukan nilai VIF:

a. Hipotesis:

H_0 : `Terdapat multikolinieritas

H_1 : Tidak terdapat multikolinieritas

b. Statistik uji VIF:

$$VIF = \frac{1}{1 - R^2} \quad (2.9)$$

c. Daerah kritis:

Apabila $VIF > 10$ maka terindikasi terdapat multikolinieritas, begitupun sebaliknya, apabila nilai $VIF \leq 10$ maka tidak terindikasi terdapat multikolinieritas (Yudiatmaja, 2013).

2.6 K-Means

K-means adalah salah satu metode non-hierarki yang mempartisi data yang ada menjadi beberapa *cluster*. Data dengan karakteristik yang sama dikumpulkan dalam kelompok yang sama juga, sedangkan data dengan karakteristik berbeda dikumpulkan pada kelompok yang lain. Algoritma *clustering k-means* memiliki sifat iteratif, numerik dan masuk dalam kategori *unsupervised learning*. Algoritma ini diusulkan pertama kali oleh MacQueen kemudian dikembangkan oleh

Hartigan dan Wong. Berikut adalah algoritma *k-means clustering* (Wahyudi, 2020):

1. Menentukan jumlah data yang akan di *cluster*
2. Menetapkan jumlah *cluster* (K) yang akan dibentuk.
3. Menentukan titik pusat *cluster* (*centroid*) awal secara *random*.

Penentuan *centroid* awal dilakukan secara acak (*random*) dari objek-objek yang tersedia sebanyak K *cluster*. Kemudian, untuk menghitung *centroid cluster* ke- i berikutnya dengan menggunakan persamaan yang akan ditampilkan pada tahap selanjutnya.

4. Menghitung jarak setiap objek terhadap tiap *centroid* dari masing-masing *cluster*.

Jarak antara objek dengan *centroid* dari masing-masing *cluster* dapat menggunakan jarak euclidean (*euclidean distance*) pada persamaan (2.1) hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*.

5. Menghitung nilai *centroid* baru dengan keanggotaan *cluster* yang terbentuk.

Nilai *centroid* baru diperoleh dari rata-rata (*mean*) *cluster* yang bersangkutan dengan menggunakan persamaan sebagai berikut:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad (2.10)$$

Keterangan:

V_{ij} = *Centroid cluster* ke- i untuk variabel ke- j

N_i = Jumlah anggota *cluster* ke- i

X_{kj} = Nilai data ke- k variabel ke- j untuk *cluster* tersebut

$i.k$ = Indeks dari *cluster*

j = Indeks variabel

6. Menghitung kembali setiap objek menggunakan *centroid* baru hingga tidak ada perubahan pada anggota masing-masing *cluster*.

Jika anggota *cluster* sudah tidak ada perubahan, maka proses *clustering* dinyatakan selesai. Namun, apabila anggota *cluster* masih mengalami perubahan maka kembali diulang kembali langkah ke-4 dan ke-5 hingga jumlah iterasi terpenuhi (anggota *cluster* tidak mengalami perubahan).

2.7 Metode *Elbow*

Metode *elbow* adalah salah satu metode yang bisa digunakan untuk menentukan jumlah *cluster* paling *optimal*. Metode *elbow* bertujuan untuk menentukan jumlah *cluster optimal* dengan memberikan informasi berupa persentase hasil dari perbandingan antar jumlah *cluster* yang akan membentuk siku pada suatu titik. Hasil persentase yang berbeda dari setiap nilai *cluster* tersebut ditunjukkan dengan menggunakan grafik sebagai sumber informasinya (Harani, 2020).

Pada metode *elbow* dilakukan perhitungan nilai selisih penurunan nilai *Within Cluster Sum Of Square* (WCSS) atau biasa disebut dengan *Sum of Square Error* (SSE) yang paling besar dan berbentuk siku. SSE merupakan jumlah kuadrat jarak setiap titik data dengan *centroid* *clusternya*, semakin besar jumlah *cluster* k maka nilai SSE akan semakin kecil. Berikut persamaan SSE (*Sum of Square Error*) (Harani, 2020):

$$SSE = \sum_{k=1}^k \sum_{x_i \in S_k} \|x_i - c_k\|^2 \quad (2.11)$$

Keterangan:

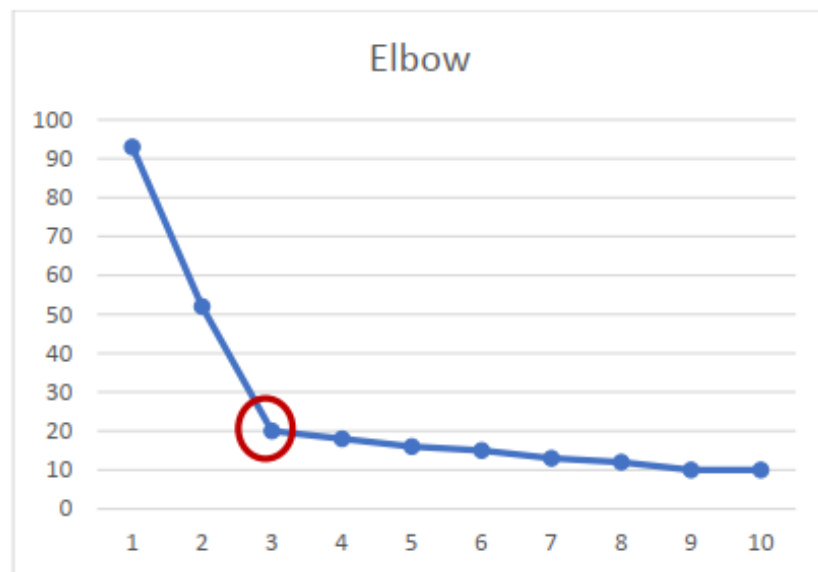
K = Jumlah *cluster*

x_i = Data ke-i

c_k = *Centroid cluster*

S_K = K *cluster* yang terbentuk

Setelah dilakukan perhitungan nilai SSE, nilai *cluster* k dapat mengalami perubahan (penurunan). Nilai k akan turun secara bertahap hingga hasil dari nilai K tersebut memiliki nilai yang stabil. Berikut adalah contoh grafik pada metode *elbow*:



Gambar 2.1 Contoh Grafik Metode *Elbow*

Pada grafik tersebut menunjukkan nilai K (jumlah *cluster*) yang ada pada sumbu x dan nilai SSE sebagai tolak ukur kesalahan pada sumbu y. Terlihat bahwa pada grafik Gambar 2.1, *error* menurun drastis ketika K berubah dari 2

menuju ke-3, namun setelah itu *error* menurun perlahan setelah melalui nilai $K = 3$, sehingga grafik terlihat hampir berbentuk ‘siku’ pada nilai $K = 3$. Dengan demikian $K = 3$ adalah nilai *optimal* dari K .

2.8 COVID-19

Coronavirus Disease (COVID-19) adalah penyakit yang disebabkan oleh *Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2)*. Selain menginfeksi manusia, virus ini juga dapat keluarga besar kelompok coronavirus tersebut juga dapat menyerang hewan. Saat menyerang manusia, virus corona biasanya memberikan dampak infeksi pernapasan ringan maupun sedang misalnya *influenza*, namun virus tersebut dapat juga memberikan dampak berat pada pernapasan seperti *Pneumonia*, *Middle East Respiratory Syndrome (MERS)* dan *Severe Acute Respiratory Syndrome (SARS)*. Metode utama penularan penyakit ini adalah melalui tetesan (*droplet*) yang dikeluarkan ketika seseorang batuk ataupun bersin. Coronavirus sendiri merupakan jenis infeksi lain yang ditemukan di Wuhan, Hubei, China menjelang akhir tahun 2019. Karenanya, infeksi tersebut diberi nama *Coronavirus Disease 2019* yang disingkat menjadi COVID-19. Sejak COVID-19 ditemukan, sampai kini virus tersebut terus menyebar dan menjadi pandemi diseluruh dunia (Setiawan, 2020).

Penyakit *Coronavirus* dapat menyebabkan dampak ringan, sedang, hingga kematian. Respon tubuh apabila terkena infeksi virus tersebut adalah suhu tubuh lebih dari 38 derajat celcius, batuk dan sesak. Gejala ringan yang terjadi karena infeksi COVID-19 antara lain batuk, demam, dan sesak napas. Adapun gejala yang paling ringan tidak menentu, sedangkan gejala beratnya adalah susah

bernapas, takipnea (frekuensi bernapas lebih dari 30 kali per menit pada orang dewasa), saturasi oksige kurang dari 90% udara luar (Burhan & dkk, 2020). Adapun cara yang dapat dilakukan untuk mendeteksi infeksi virus COVID-19 pada diri seseorang dilakukakannya pemeriksaan berikut (Karyono, Indriyani, & Rohadin, 2020):

1. Rapid tes untuk membedakan antibodi (IgM dan IgG) yang dikirim oleh tubuh untuk melawan infeksi COVID-19.
2. Tes swab atau tes PCR (*Polymerase Chain Reaction*) untuk mengenali infeksi virus corona yang ada di dahak.
3. CT *scan/rongten* dada yang fungsi mendeteksi infiltrat cairang yang ada di paru-paru (*in filtrat*).

Istilah-istilah yang berhubungan dengan COVID-19, antara lain:

1. *Suspect*

Seseorang disebut *suspect* apabila dalam dirinya mengalami paling tidak satu dari kriteria dibawah ini:

- a. Mengalami gejala Infeksi Saluran Pernafasan (ISPA), misalnya demam atau pernah mengalami demam dengan suhu diatas $38^{\circ}C$ dan salah satu gejala penyakit yang berkaitan dengan pernapasan seperti asma, batuk dan lain-lain.
- b. Pernah kontak dengan orang yang kategori *probable* atau justru sudah terkonfirmasi COVID-19 dalam 14 hari terakhir.

2. *Probable*

Seseorang yang masih dalam kategori suspect dan memiliki gejala ISPA berat, gagal nafas, atau bahkan meninggal dunia. Namun belum ada hasil pemeriksaan yang memastikan bahwa dirinya positif COVID-19.

3. Konfirmasi atau Positif

Kondisi ketika berdasarkan hasil pemeriksaan tes laboratorium seseorang dinyatakan positif terinfeksi virus COVID-19.

4. Kematian/Meninggal

Termasuk kategori ini ketika kondisi seseorang yang termasuk kategori probable atau telah dinyatakan positif COVID-19 meninggal dunia.

2.9 Analisis *Cluster* dalam Islam

Tujuan dari proses analisis *clustering* adalah pengelompokan data atau objek. Sehingga objek yang memiliki karakteristik semirip mungkin, akan dikelompokkan pada *cluster* yang sama dan objek yang tidak ada kemiripan karakteristik akan terkelompokkan pada *cluster* yang lain (Han & Kamber, 2001). Dalam Al-Qur'an, Allah memberikan banyak informasi untuk manusia melalui berbagai ayat sebagai petunjuk menuju jalan yang lurus, yakni jalan menuju kebahagiaan dunia dan akhirat. Salah satu petunjuk yang sesuai dengan konsep analisis *clustering* yakni pengelompokan manusia pada *yaumulakhir*, sebagaimana yang termaktub pada QS. Al Waqiah ayat 7-10 berbunyi:

وَكُنْتُمْ أَزْوَاجًا ثَلَاثَةً (٧) فَأَصْحَابُ الْمَيْمَنَةِ مَا أَصْحَابُ الْمَيْمَنَةِ (٨) وَأَصْحَابُ الْمَشْأَمَةِ مَا أَصْحَابُ الْمَشْأَمَةِ

(٩) وَالسَّابِقُونَ السَّابِقُونَ (١٠)

Artinya:

“Dan kamu menjadi tiga golongan. Yaitu golongan kanan, alangkah mulianya golongan kanan itu. dan golongan kiri, alangkah sengsaranya golongan kiri itu.

Dan orang-orang yang paling dahulu(beriman), merekalah yang paling dahulu (masuk surga)”

Ayat tersebut merupakan salah satu contoh pegelompokkan pada firman Allah yang memberitakan bahwa pada hari kiamat nantinya manusia akan dibagi 3 kelompok sesuai dengan amal yang dikerjakan selama di dunia. Menurut Tafsir Ibnu Katsir (Ghoffar & Al-Atsari, 2004) kelompok-kelompok tersebut yakni:

1. *Ashab Al-Yamin* (Golongan Kanan)

Ashab Al-Yamin (golongan kanan) adalah kelompok manusia yang teguh iman kepada Allah, taat kepada perintah-perintah Allah dan menjauhi larangan-larangan Allah. *Ashab Al-Yamin* ini nantinya akan mengambil dari tangan kanan catatan amal mereka, kemudian diarahkan kekanan menuju surga. Mengambil dari tangan kanan disini menunjukkan keadaan yang sangat baik dan menyenangkan.

2. *Ashab Al Syimal* (Golongan Kiri)

Ashab Al Syimal (golongan kiri) adalah kelompok manusia yang sering ingkar dan berbuat kemaksiatan. Golongan kiri ini nantinya akan menerima buku catatan amal mereka dengan tangan kiri, yang menandakan penghuni neraka dan akan mendapatkan hukuman atas perbuatan yang telah dilakukan.

3. *Al Sabiqun* (Golongan Tengah)

Al Sabiqun yakni kelompok orang-orang yang beriman terdahulu pada dakwah rosulullah. Golongan *al sabiqun* ini merupakan orang-orang yang paling khusus, terhormat dan lebih dekat dibandingkan orang golongan kanan sebagai pemuka mereka semua. Karena diantara mereka adalah golongan rasul, golongan nabi, golongan orang yang jujur/benar(*as-siddiqun*) dan syuhada yang jumlahnya tidak lebih banyak dari ashabulyamin.

QS.Al-Waqi'ah ayat 7-10 diatas adalah salah satu contoh konsep analisis *cluster* yang dikelompokkan menjadi tiga *cluster*, namun hal ini bukan berarti bahwa analisis *cluster* harus selalu dikelompokkan menjadi tiga kelompok saja, pada topik bahasan di ayat lainnya seperti pada QS. Az-Zumar ayat 73:

وَسِيقَ الَّذِينَ اتَّقَوْا رَبَّهُمْ إِلَى الْجَنَّةِ زُمَرًا ... (٧٣)

Artinya: “Dan orang-orang yang bertakwa kepada Tuhan dibawa ke dalam surga berombong-rombongan (pula),...”.

Menurut Tafsir Ibn Katsir (Ghoffar & Al-Atsari, 2004) pada ayat tersebut dijelaskan bahwa orang-orang yang bertaqwa akan dimuliakan dan diberikan penghormatan dengan cara dibimbing menuju surga juga secara berkelompok. Mereka yang bertaqwa diantar ke surga dan dibukakan pintu surga dengan penuh pemuliaan. Masing-masing kelompok bersama dengan kelompok yang sesuai derajat amalnya. Sehingga setiap kelompok tersebut memiliki kriteria masing-masing. Misal, kelompok nabi bersama nabi, orang yang jujur bersama dengan yang serupa, syuhada bersama syuhada, dan seterusnya. Hal tersebut sesuai dengan konsep analisis *k-means cluster*, yang berusaha untuk mempartisi data yang tersedia kedalam kelompok dengan sifat yang relatif homogen. Objek-objek yang sama atau mirip kedekatan jaraknya antar variabel, akan ditempatkan di *cluster* yang sama, sedangkan objek-objek lain yang tidak sejenis akan dikelompokkan dalam *cluster* yang berbeda, sehingga diperoleh kelompok wilayah yang memiliki homogenitas dan heterogenitas antar kelompok yang tinggi.

2.10 Analisis Cluster COVID-19 di Provinsi Jawa Timur

Analisis *cluster* adalah salah satu analisis yang bertujuan untuk mengelompokkan objek-objek yang memiliki kemiripan dalam satu kelompok (*cluster*). Sebelum dilakukan analisis *cluster*, objek yang akan dikelompokkan haruslah memenuhi uji asumsi terlebih dahulu. Uji asumsi yang harus terpenuhi yakni uji multikolinieritas. Uji multikolinieritas terjadi apabila VIF lebih dari 10. Apabila data sudah memenuhi kedua uji asumsi, maka langkah selanjutnya dilakukan analisis *cluster*. Jika terjadi multikolinieritas, artinya terdapat hubungan yang linier antar variabel satu dengan lainnya. Oleh karena itu, suatu data haruslah tidak terdapat multikolinieritas. Apabila ketiga uji asumsi sudah terpenuhi, maka langkah selanjutnya yakni melakukan analisis *cluster*. Sebelum dilakukan pengelompokan, perlu dilakukan standarisasi data terlebih dahulu supaya data yang diolah memiliki rentang yang tidak jauh berbeda.

BAB III

METODOLOGI PENELITIAN

3.1 Jenis Penelitian

Jenis penelitian ini adalah penelitian deskriptif kuantitatif dan studi literatur. Dikatakan deskriptif kuantitatif karena data yang digunakan pada penelitian ini merupakan data yang sudah ada dalam bentuk numerik (data numerik) yang selanjutnya data tersebut dianalisis serta disusun berdasarkan kebutuhan dan aturan penelitian. Deskriptif data terlebih dahulu dilakukan sebelum analisis *cluster* kasus COVID-19 di Provinsi Jawa Timur guna untuk melihat gambaran umum dari tiap variabel data.

3.2 Data dan Sumber Data

Data penelitian ini menggunakan data sekunder dari kasus pandemi COVID-19 di Provinsi Jawa Timur periode 31 Juli 2021 yang diperoleh dari laman resmi Pemerintah Provinsi Jawa Timur untuk menginformasikan *update* penyebaran COVID-19 di Jawa Timur. Adapun data yang digunakan yakni data jumlah kasus COVID-19 dari 38 kabupaten dan kota di Provinsi Jawa Timur Lampiran 1. Jumlah kasus COVID-19 dalam penelitian ini yakni banyaknya pasien yang sembuh dan pasien yang meninggal akibat COVID-19.

3.3 Lokasi Penelitian

Penelitian ini dilakukan dikediaman peneliti yang berdomisili di Kelurahan Bandulan Kecamatan Sukun Kota Malang Provinsi Jawa Timur. Hal ini

dikarenakan data penelitian yang digunakan merupakan data sekunder yang dapat diakses sendiri melalui laman Pemerintah Provinsi Jawa Timur.

3.4 Teknik Pengumpulan Data

Pengumpulan data pada penelitian ini dengan mengakses internet pada laman Jatim Tanggap COVID-19. Jatim Tanggap COVID-19 merupakan laman resmi Pemerintah Provinsi Jawa Timur yang yang dikhususkan untuk menginformasikan *update* setiap hari penyebaran COVID-19 di Provinsi Jawa Timur. Pengumpulan data dilakukan pada 05 Agustus 2021 pukul 21.18 WIB sesuai dengan kebutuhan peneliti.

3.5 Instrumen Penelitian

Instrumen penelitian merupakan alat yang digunakan untuk mengumpulkan data. Sehingga pada penelitian ini instrumennya adalah peneliti sendiri. Peneliti berperan dalam mengumpulkan bahan literatur yang akan digunakan, mengumpulkan data, menentukan variabel yang digunakan serta menganalisis hingga menjadi laporan hasil penelitian. Penelitian ini menggunakan dua variabel data sebagai berikut:

X1 : Kasus pasien sembuh akibat COVID-19 (orang)

X2 : Kasus pasien meninggal akibat COVID-19 (orang)

Penelitian ini juga menggunakan bantuan beberapa *software* yakni:

1. *Software python* untuk mengetahui grafik *elbow* beserta nilai SSE dan juga hasil pengelompokan dari jumlah *cluster* yang telah ditentukan.

2. *Software minitab* untuk melakukan deteksi *outlier*, mengetahui statistika deskriptif, memunculkan *scatterplot* antara variabel.
3. *Software microsoft excel* untuk mengkonfirmasi proses normalisasi data, uji multikolinieritas dan proses *k-means*.

3.6 Teknik Analisis Data

Teknik analisis data merupakan langkah yang akan paling menentukan terhadap hasil suatu penelitian, oleh karenanya teknik analisis data memiliki peran yang sangat penting dalam sebuah penelitian. Adapun langkah-langkah yang dapat dilakukan untuk menganalisis data kasus pasien sembuh dan meninggal akibat COVID-19 di Jawa Timur adalah sebagai berikut:

1. Mendeskripsikan data

Sebelum dilakukan pengelompokan data, perlu dilakukan deskripsi data terlebih dahulu untuk mengetahui parameter distribusi dari masing-masing variabel yang digunakan.

2. Deteksi *Outlier*

Uji deteksi *outlier* perlu untuk dilakukan guna mengetahui apakah terdapat data yang mempunyai perbedaan yang signifikan antara suatu data dengan data lainnya.

3. Melakukan uji asumsi

Pada proses pengelompokan asumsi yang harus terpenuhi yakni tidak adanya multikolinieritas. Sedangkan untuk mendeteksi adanya multukolinieritas dapat dilakukan dengan cara menghitung nilai VIF.

4. Melakukan normalisasi data

Normalisasi data dilakukan apabila data memiliki rentang atau skala yang berbeda, sehingga terdapat data yang terlalu tinggi dan terlalu rendah. Karena data pada penelitian ini memiliki kesamaan satuan namun terdapat perbedaan rentang data, sehingga proses normalisasi dilakukan untuk membawa semua data pada rentang atau skala yang sama yakni 0-1.

5. Memilih ukuran jarak

Pengukuran jarak pada pengelompokan juga perlu untuk dilakukan. Pada penelitian ini jarak yang digunakan adalah *euclidean distance*.

6. Menentukan jumlah *cluster*

Pada penelitian ini menggunakan pengelompokan non-hierarki, sehingga peneliti dapat menentukan terlebih dahulu jumlah *cluster* (K) yang akan terbentuk. Sehingga untuk memperoleh K yang optimal, peneliti menggunakan metode *elbow* dengan melihat grafik yang terbentuk.

7. Melakukan pengelompokan *k-means* dan interpretasi output analisis *k-means cluster*

Pada proses pengelompokan, data yang dianalisis dengan menggunakan bantuan *software python* dan *microsoft excel*. Selanjutnya *output* diinterpretasi untuk membaca hasil analisis yang telah dilakukan serta memberikann nama pada setiap kelompok yang terbentuk.

8. Menentukan ciri kelompok

Untuk menggambarkan ciri kelompok yang terbentuk, peneliti menggunakan scatter plot untuk mengetahui relasi tiap variabel data dari kelompok yang terbentuk.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Statistika Deskriptif

Statistika deskriptif dilakukan untuk memberikan informasi mengenai gugus data secara umum mengenai karakteristik variabel yang digunakan peneliti dan agar lebih mudah dipahami, maka berikut disediakan Tabel 4.1 pada kasus COVID-19 di Provinsi Jawa Timur.

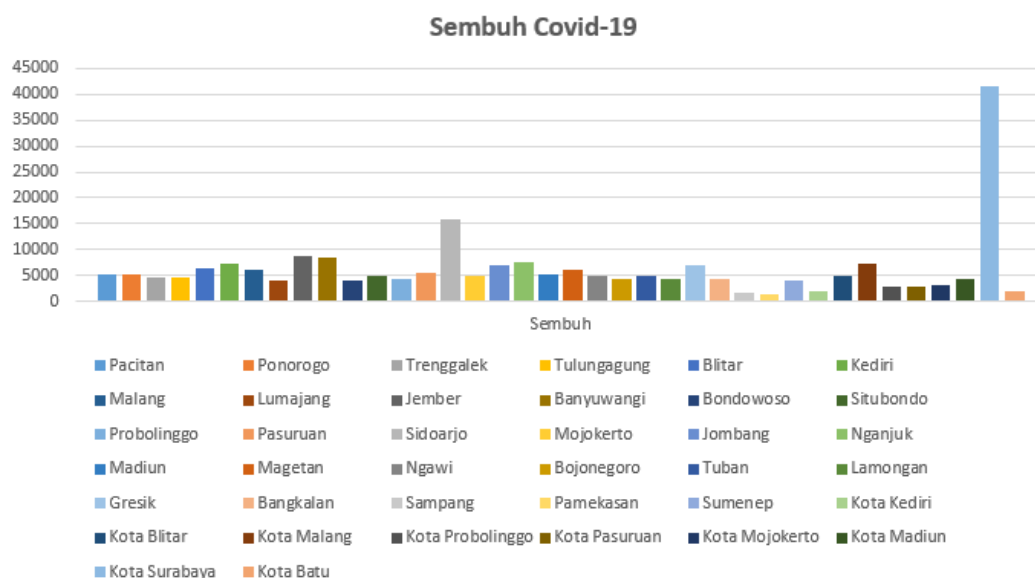
Tabel 4.1 Statistika Deskriptif

Variabel	N	Mean	St.Deviasi	Minimum	Maksimum
Sembuh (X1)	38	6140	6390	1488	41448
Meninggal (X2)	38	5440	374.7	110	1888

(Sumber: Output Minitab)

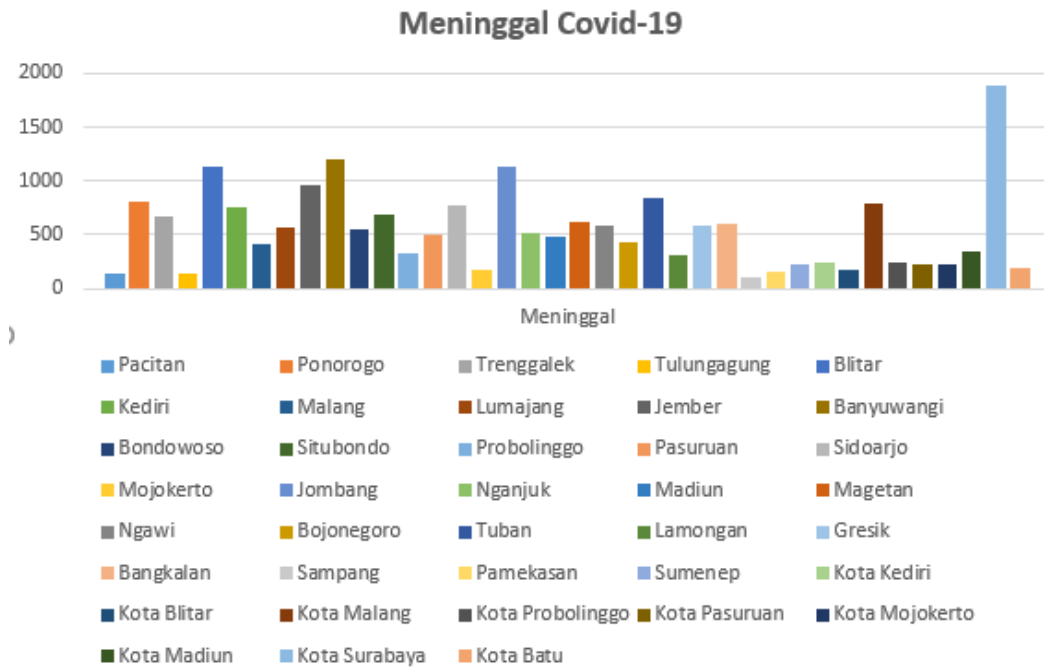
Berdasarkan Tabel 4.1 tersebut menjelaskan statistika deskriptif dari variabel yang terdapat pada data kasus COVID-19 Provinsi Jawa Timur yang digunakan peneliti. Pada kolom N, kedua variabel memiliki 38 data, artinya dapat diketahui bahwa Provinsi Jawa Timur memiliki 38 kabupaten dan kota. Pada variabel pasien sembuh akibat COVID-19 bulan periode 31 Juli 2021 (X1), paling rendah (minimum) ada sebanyak 1488 kasus, sedangkan kasus tertingginya (maksimal) sebanyak 41448 kasus. Sedangkan pada variabel pasien meninggal akibat COVID-19 bulan periode 31 Juli 2021 (X2), paling rendah (minimum) ada

sebanyak 110 kasus yakni Kabupaten Pamekasan, sedangkan kasus tertingginya (maksimal) sebanyak 1888 kasus yakni Kota Surabaya. Kasus sembuh dari COVID-19 memiliki rata-rata sebanyak 6140 kasus sembuh dan sebanyak 5440 kasus meninggal di seluruh Provinsi Jawa Timur. Pada variabel meninggal akibat COVID-19 Kasus meninggal dari COVID-19 memiliki rata-rata lebih kecil daripada kasus sembuh, hal tersebut menunjukkan bahwa pasien sembuh dari COVID-19 lebih banyak daripada pasien yang meninggal akibat COVID-19 di Provinsi Jawa Timur. Kemudian dilihat dari nilai standar deviasi, jumlah kasus sembuh lebih tinggi yakni 6390 pasien sembuh, sedangkan yang meninggal akibat COVID-19 yakni 374.7 pasien. Hal tersebut dapat diartikan bahwa jumlah pasien sembuh dan meninggal tiap kabupaten-kota di Provinsi Jawa Timur tidak merata atau terdapat kesenjangan pasien antar kota-kabupaten di Provinsi Jawa Timur.



Gambar 4.1 Grafik Penyebaran Kasus Sembuh COVID-19 di Provinsi Jawa Timur

Berdasarkan Gambar 4.1 dapat dilihat bahwa variabel X1 (sembuh akibat COVID-19) pada data kasus penyebaran COVID-19 di Provinsi Jawa Timur memiliki angka tertinggi pada kota Surabaya yakni 41448 pasien sembuh dan memiliki angka terendah pada Kabupaten Pamekasan yakni 1488 pasien sembuh.

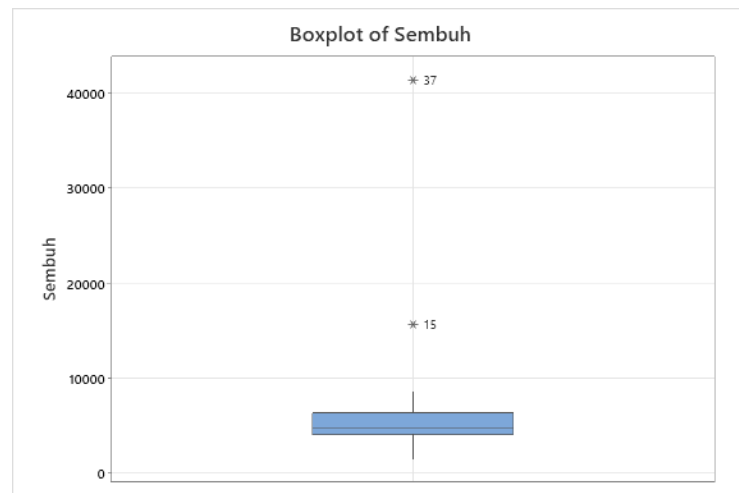


Gambar 4.2 Grafik Penyebaran Kasus Meninggal Akibat COVID-19 di Provinsi Jawa Timur

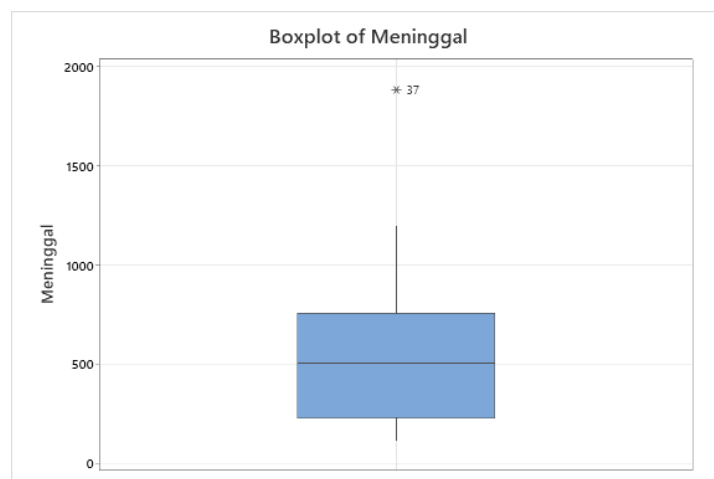
Berdasarkan Gambar 4.2 dapat dilihat bahwa variabel X2 (meninggal akibat COVID-19) pada data kasus penyebaran COVID-19 di Provinsi Jawa Timur memiliki angka tertinggi pada Kota Surabaya yakni 1888 pasien meninggal dan memiliki angka terendah pada Kabupaten Sampang yakni 110 pasien meninggal. Dari beberapa wilayah Kabupaten-Kota Provinsi Jawa Timur, dapat diketahui bahwa wilayah tertinggi kasus sembuh sama dengan wilayah kasus meninggal akibat COVID-19 dan Kota Surabaya merupakan wilayah yang mendominasi.

4.2 Deteksi *Outlier*

Sebelum melakukan langkah selanjutnya yakni uji asumsi analisis *cluster*, perlu dilakukan deteksi data terhadap *outlier*. Jika terdapat *outlier* artinya data memiliki perbedaan yang signifikan terhadap data lainnya. Berikut merupakan hasil deteksi *outlier* menggunakan boxplot.



Gambar 4.3 Boxplot X1
(Sumber: Output Minitab)



Gambar 4.4 Boxplot X2
(Sumber: Output Minitab)

Gambar 4.3 dan Gambar 4.4 merupakan output pendeteksian *outlier* dengan menggunakan *software* minitab. Dapat dilihat pada Gambar 4.3 dan Gambar 4.4 menunjukkan bahwa adanya *outlier* pada variabel X1 dan X2. Pada variabel X1 terdapat *outlier* pada urutan kabupaten/kota ke-15 dan ke-37 yakni Kabupaten Sidoarjo dan Kota Surabaya, sedangkan pada variabel X2 terdapat *outlier* pada urutan kabupaten-kota ke-37 yakni kota Surabaya. Hal tersebut ditunjukkan dengan adanya titik diluar *boxplot*. Terdeteksinya *outlier* pada variabel X1 dikarenakan kabupaten-kota ke-15 dan ke-37 memiliki banyak kasus sembuh dari COVID-19 yang tinggi dibandingkan dengan yang lainnya. Pada *boxplot* dari variabel X2 juga menunjukkan adanya *outlier* karena kabupaten/kota ke-37 memiliki banyak pasien yang meninggal akibat COVID-19 yang tinggi dibandingkan dengan yang lainnya. Selanjutnya data outlier tersebut tidak dihapuskan dan tetap digunakan dalam analisis penelitian ini karena variabel yang dipertimbangkan cukup memadai dan dapat mewakili populasinya.

4.3 Uji Asumsi Analisis *Cluster*

Uji Multikolinieritas

Uji multikolinieritas dilakukan untuk menguji adanya variabel independen yang memiliki kesamaan karakteristik antar variabel independen lain. Pada penelitian ini cara yang digunakan untuk mengetahui adanya indikasi multikolinieritas adalah dengan mencari nilai *Variance Inflation Factor* (VIF) menggunakan persamaan (2.9). Sebelum mencari nilai VIF untuk X_2 pada langkah ini, peneliti terlebih dahulu mencari nilai matriks H dengan persamaan (2.5):

$$\begin{aligned}
H &= \begin{bmatrix} 1 & 5095 \\ 1 & 5201 \\ 1 & 4739 \\ \vdots & \vdots \\ 1 & 41448 \\ 1 & 1873 \end{bmatrix} \left(\begin{bmatrix} 1 & \dots & 1 \\ 5095 & \dots & 1873 \end{bmatrix} \begin{bmatrix} 1 & 5095 \\ 1 & 5201 \\ 1 & 4739 \\ \vdots & \vdots \\ 1 & 41448 \\ 1 & 1873 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & \dots & 1 \\ 5095 & \dots & 1873 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 5095 \\ 1 & 5201 \\ 1 & 4739 \\ \vdots & \vdots \\ 1 & 41448 \\ 1 & 1873 \end{bmatrix} \begin{bmatrix} 38 & 233312 \\ 233312 & 2943077836 \end{bmatrix}^{-1} \begin{bmatrix} 1 & \dots & 1 \\ 5095 & \dots & 1873 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 5095 \\ 1 & 5201 \\ 1 & 4739 \\ \vdots & \vdots \\ 1 & 41448 \\ 1 & 1873 \end{bmatrix} \begin{bmatrix} 0.051270928 & -4.06449E-06 \\ -4.06449E-06 & 6.61992E-10 \end{bmatrix} \begin{bmatrix} 1 & \dots & 1 \\ 5095 & \dots & 1873 \end{bmatrix}
\end{aligned}$$

Sehingga diperoleh matriks H dengan ordo 38x38 pada Lampiran 2. Kemudian dilanjutkan dengan perhitungan nilai jumlah kuadrat regresi (SS_R) menggunakan persamaan (2.6):

$$\begin{aligned}
SS_R &= y' \left[H - \left(\frac{1}{n} \right) J \right] y \\
&= [146 \quad \dots \quad 188] \left(\begin{bmatrix} 0.2704 & \dots & 0.0292 \\ \vdots & \vdots & \vdots \\ 0.0292 & \dots & 0.0384 \end{bmatrix} - \left(\frac{1}{38} \right) \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \right) \begin{bmatrix} 146 \\ \vdots \\ 188 \end{bmatrix}
\end{aligned}$$

dan jumlah kuadrat total (SS_T) dengan persamaan (2.7):

$$\begin{aligned}
SS_T &= y' \left[I - \left(\frac{1}{n} \right) J \right] y \\
&= [146 \quad \dots \quad 188] \left(\begin{bmatrix} 1 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & 1 \end{bmatrix} - \left(\frac{1}{38} \right) \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \vdots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \right) \begin{bmatrix} 146 \\ \vdots \\ 188 \end{bmatrix}
\end{aligned}$$

sehingga didapat:

Tabel 4.2 Hasil Perhitungan SS_T dan SS_R

SS_T	5194854.974
SS_R	2872852.14

Selanjutnya, mencari nilai koefisien determinasi (R^2) dengan rumus pada persamaan (2.8) diperoleh:

$$\begin{aligned} R^2 &= \frac{5194854.974}{2872852.14} \\ &= 0.553018738 \end{aligned}$$

Selanjutnya menghitung nilai VIF dengan rumus pada persamaan (2.9) diperoleh:

$$\begin{aligned} VIF &= \frac{1}{1 - 0.553018738^2} \\ &= 2.237230247 \end{aligned}$$

1. Hipotesis masalah

H_0 : Terdapat multikolinieritas

H_1 : Tidak terdapat multikolinieritas

2. Statistik Uji VIF

$$VIF = 2.237230247$$

3. Daerah Kritis

Tolak H_0 jika nilai $VIF \leq 10$

$$\text{Nilai } VIF \leq 2.23723024710$$

4. Keputusan

Tolak H_0

5. Kesimpulan

Tidak terdapat multikolinieritas

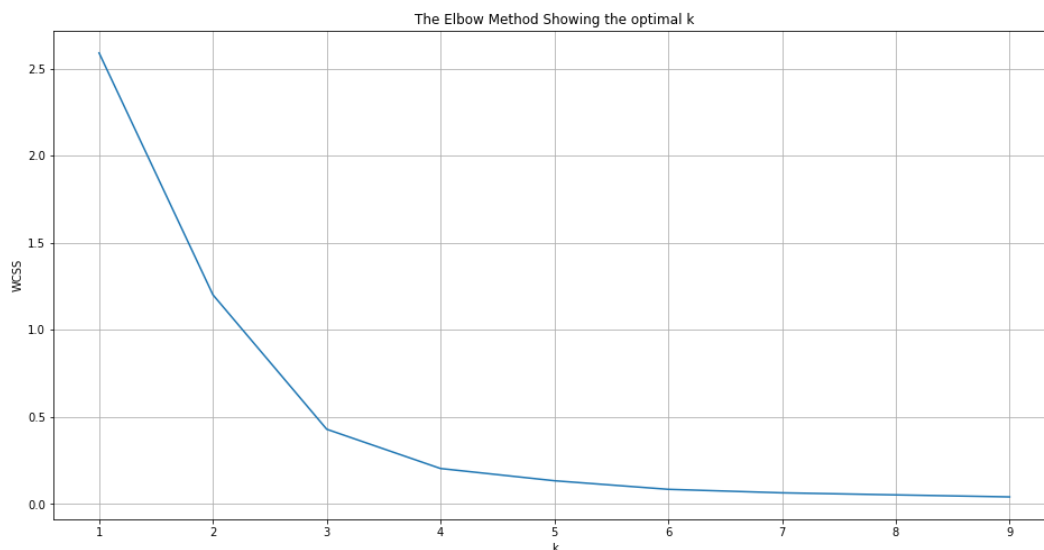
Karena nilai $VIF \leq 10$, maka dapat disimpulkan bahwa tidak terjadi multikolinieritas, sehingga dapat dilakukan langkah penelitian selanjutnya.

4.4 Normalisasi Data

Suatu data perlu dilakukan standarisasi atau normalisasi data apabila data yang digunakan dalam suatu penelitian memiliki satuan yang berbeda, selain itu rentang data yang jauh berbeda juga mempengaruhi proses pengelompokan. Data yang digunakan pada penelitian ini merupakan data dengan satuan yang sama, namun memiliki skala data yang jauh berbeda, sehingga perlu dilakukan normalisasi data untuk mempersempit rentang data supaya tidak ada data yang memiliki skala yang sangat tinggi dan sangat rendah. Metode normalisasi yang digunakan pada penelitian ini adalah *Min Max Normalization*. Pada penelitian ini semua data dinormalkan pada skala 0 sampai 1 dengan menggunakan rumus pada persamaan (2.2). Hasil normalisasi data selengkapnya dapat dilihat pada Lampiran 5.

4.5 Penentuan Jumlah Kluster

Setelah uji asumsi terpenuhi, langkah selanjutnya yakni menentukan jumlah *cluster (K) optimal* dengan menggunakan metode *elbow* yang dilakukan menggunakan bantuan *software python*. Berdasarkan *script* pada Lampiran 14, diperoleh *output* berupa grafik metode *elbow* sebagai berikut:



Gambar 4.5 Grafik Metode *Elbow*
(Sumber: Output python)

Tabel 4.3 Tabel Nilai SSE

K	SSE	Selisih
1	2.589284359	-
2	1.200641089	1.38864327
3	0.428532217	0.772108872
4	0.20296737	0.225564847
5	0.132813484	0.070153886
6	0.083551716	0.049261768
7	0.062698493	0.020853223
8	0.05280624	0.009892253
9	0.039960075	0.012846165

Pada Gambar 4.5 grafik terlihat bahwa penurunan yang paling signifikan terletak pada $K = 2$ dengan nilai selisih SSE adalah 1.38864327, namun pada penurunan tersebut tidak *optimal* dikarenakan tidak berbentuk siku dan sesudah nilai $K = 2$ masih ada penurunan grafik yang signifikan dan berbentuk siku. Dengan melihat Tabel 4.3, K *optimal* pada penelitian ini adalah $K = 4$ dengan nilai SSE adalah 0.20296737. Hal tersebut dikarenakan terjadi penurunan grafik yang signifikan dengan nilai selisih SSE yakni 0.225564847 dan dapat dikatakan berbentuk siku

karena grafik mengalami penurunan secara bertahap artinya nilai selisih SSE pada tiap K sesudah $K = 4$ memiliki nilai yang stabil.

4.6 Pengelompokan dan Interpretasinya

Langkah selanjutnya, data kasus COVID-19 Provinsi Jawa Timur dilakukan pengelompokan atau *clustering* dengan menggunakan *k-means* sebagai berikut:

1. Menginputkan variabel data

Data yang diinputkan pada proses pengelompokan ini merupakan data yang sudah dinormalisasi pada Lampiran 5. Variabel yang digunakan yakni X_1 (kasus pasien sembuh) dan X_2 (kasus pasien meninggal) dari 38 kabupaten-kota di Provinsi Jawa Timur.

2. Menentukan titik pusat *cluster* (*centroid*) awal secara random.

Titik pusat awal pada proses ini dipilih secara acak (*random*) dengan jumlah *centroid* awal *cluster* sebanyak 4 pada masing-masing variabel. Berikut adalah *centroid* awal yang digunakan:

Tabel 4.4 *Centroid* Awal

<i>Centroid</i> Awal	X1	X2
<i>C0</i>	1	1
<i>C1</i>	0.11957	0.173791
<i>C2</i>	0.05976	0.036558
<i>C3</i>	0	0

3. Menghitung jarak setiap data terhadap *centroid* dari masing-masing *cluster*.

Perhitungan jarak pada penelitian ini dilakukan dengan menggunakan *euclidean distance* pada persamaan (2.1). Berikut adalah contoh perhitungan

euclidean distance pada data ke-1 (Kabupaten Pacitan) terhadap masing-masing *centroid*:

$$\begin{aligned}d_{(0.0)} &= \sqrt{(0.090265265 - 1)^2 + (0.020247469 - 1)^2} \\ &= 1.336986\end{aligned}$$

$$\begin{aligned}d_{(0.1)} &= \sqrt{(0.09026527 - 0.11956957)^2 + (0.02024747 - 0.173790776)^2} \\ &= 0.156315\end{aligned}$$

$$\begin{aligned}d_{(0.2)} &= \sqrt{(0.090265265 - 0.05975976)^2 + (0.020247469 - 0.3655793)^2} \\ &= 0.034592\end{aligned}$$

$$\begin{aligned}d_{(0.3)} &= \sqrt{(0.090265265 - 0)^2 + (0.020247469 - 0)^2} \\ &= 0.092508\end{aligned}$$

Hasil perhitungan *euclidean distance* selengkapnya tertera pada Lampiran 6 sampai Lampiran 13. Setelah menghitung jarak setiap data menggunakan *euclidean distance*, kemudian dicari jarak minimum pada C0,C1,C2,C3 sehingga diperoleh hasil pengelompokan iterasi-1 yang tertera pada Lampiran 6.

4. Menghitung nilai *centroid* baru dengan keanggotaan *cluster* yang terbentuk dengan menggunakan *centroid* awal.

Untuk memperoleh *centroid* baru, maka terlebih dahulu menentukan rata-rata (*mean*) setiap data pada *cluster* yang diikuti dengan menggunakan persamaan (2.10). Perhitungan rata-rata tersebut bertujuan untuk menentukan apakah *centroid* baru pada iterasi selanjutnya mengalami perubahan atau tidak (Lampiran 7). Setelah diketahui nilai rata-rata (*mean*) pada setiap *cluster* pada iterasi-1, selanjutnya yakni mencari nilai *centroid* baru (*centroid-2*) dengan menggunakan data hasil normalisasi dan *mean* pada setiap *cluster* pada

iterasi-1 yang diproses dengan. Sehingga diperoleh *centroid* baru yang dapat dilihat pada Tabel 4.5:

Tabel 4.5 *Centroid* Baru

<i>Centroid-2</i>	X1	X2
C0	1	1
C1	0.117540457	0.318827334
C2	0.058836337	0.055286839
C3	0.004763097	0.023809524

5. Menghitung kembali setiap objek menggunakan *centroid* baru hingga tidak ada perubahan pada anggota *cluster*.

Dari Tabel 4.9 dapat diketahui bahwa terdapat perubahan *centroid* baru pada C1,C2 dan C3 dari *centroid* awal. Sehingga diperlukan langkah perulangan dengan cara yang sama seperti sebelumnya yakni perhitungan *euclidean Distance* dengan menggunakan *centroid* baru pada Tabel 4.9. Perhitungan jarak pada iterasi ke-2 menghasilkan nilai yang tertera pada Lampiran 7. Pada pengklasteran iterasi ke-2 tersebut mengalami perubahan anggota dari pengklasteran pada iterasi-1. Oleh sebab itu proses iterasi tetap diteruskan ke iterasi-3 dan selanjutnya sampai tidak ada perubahan *centroid* dan perubahan anggota setiap *cluster* yang terbentuk. Sehingga pada proses ini iterasi dilakukan sebanyak 8 kali iterasi. Hal tersebut dikarenakan *centroid* dan anggota dari masing-masing *cluster* pada iterasi ke-8 tidak mengalami perubahan (anggotanya tetap) dari iterasi sebelumnya yakni iterasi ke-7.

Dengan menggunakan bantuan *software python* diperoleh output pengelompokan *k-means* dengan *metode elbow* sebagai berikut:

Tabel 4.6 Hasil Pengklasteran *K-Means*
(Sumber: Output python)

No	Kabupaten-Kota	Sembuh	Meninggal	Cluster
1	Pacitan	0.090265265	0.020247469	3
2	Ponorogo	0.092917918	0.38976378	1
3	Trenggalek	0.081356356	0.317210349	2
4	Tulungagung	0.074974975	0.015748031	3
5	Blitar	0.11956957	0.571428571	1
6	Kediri	0.148198198	0.362204724	1
7	Malang	0.116266266	0.173790776	2
8	Lumajang	0.06539039	0.253655793	2
9	Jember	0.17957958	0.482564679	1
10	Banyuwangi	0.177702703	0.613610799	1
11	Bondowoso	0.05975976	0.248593926	2
12	Situbondo	0.084209209	0.322834646	2
13	Probolinggo	0.074149149	0.120922385	3
14	Pasuruan	0.101101101	0.214848144	2
15	Sidoarjo	0.355905906	0.368953881	1
16	Mojokerto	0.084234234	0.03655793	3
17	Jombang	0.133983984	0.571428571	1
18	Nganjuk	0.154004004	0.230033746	2
19	Madiun	0.095845846	0.209786277	2
20	Magetan	0.112712713	0.286276715	2
21	Ngawi	0.086811812	0.264341957	2
22	Bojonegoro	0.070095095	0.181102362	2
23	Tuban	0.088213213	0.412260967	1
24	Lamongan	0.072697698	0.112485939	3
25	Gresik	0.134834835	0.268278965	2
26	Bangkalan	0.073123123	0.276715411	2
27	Sampang	0.004654655	0	3
28	Pamekasan	0	0.027559055	3
29	Sumenep	0.06458959	0.061867267	3
30	Kota Kediri	0.009884885	0.069741282	3
31	Kota Blitar	0.084884885	0.034308211	3
32	Kota Malang	0.143943944	0.383014623	1
33	Kota Probolinggo	0.030730731	0.070866142	3
34	Kota Pasuruan	0.031506507	0.066929134	3
35	Kota Mojokerto	0.044594595	0.064116985	3
36	Kota Madiun	0.071296296	0.128233971	3
37	Kota Surabaya	1	1	0
38	Kota Batu	0.009634635	0.043869516	3

Pada Tabel 4.6 menunjukkan hasil pengelompokan 4 *cluster* dari variabel kasus sembuh dan meninggal akibat COVID-19 Provinsi Jawa Timur. Berdasarkan

output tersebut, hasil pengelompokan terdapat pada kolom “*Cluster*” yang terdiri dari 4 kelompok yakni *Cluster 0*, *Cluster 1*, *Cluster 2*, dan *Cluster 3*. Selanjutnya pada tiap-tiap kelompok yang terbentuk diberikan penamaan untuk membedakan antar kelompok satu dengan kelompok lainnya. Sehingga berikut ini adalah anggota-anggota hasil pengelompokan *k-means* yang terbentuk:

Tabel 4.7 Hasil Pengelompokan *K-Means Cluster*

Penamaan <i>Cluster</i>	Kabupaten-Kota Provinsi Jawa Timur	Banyak anggota
<i>Cluster 0</i>	Kota Surabaya	1 anggota
<i>Cluster 1</i>	Kabupaten Ponorogo, Kabupaten Blitar, Kabupaten Kediri, Kabupaten Jember, Kabupaten Banyuwangi, Kabupaten Sidoarjo, Kabupaten Jombang, Kabupaten Tuban, Kota Malang	9 anggota
<i>Cluster 2</i>	Kabupaten Trenggalek, Kabupaten Malang, Kabupaten Lumajang, Kabupaten Bondowoso, Kabupaten Situbondo, Kabupaten Pasuruan, Kabupaten Nganuk, Kabupaten Madiun, Kabupaten Magetan, Kabupaten Ngawi, Kabupaten Bojonegoro, Kabupaten Gresik, Kabupaten Bangkalan	13 anggota
<i>Cluster 3</i>	Kabupaten Pacitan, Kabupaten Tulungagung, Kabupaten Probolinggo, Kabupaten Mojokerto, Kabupaten Lamongan, Kabupaten Sampang, Kabupaten Pamekasan, Kabupaten Sumenep, Kota Kediri, Kota Blitar, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Batu	15 anggota

Berdasarkan Tabel 4.7 dapat diketahui bahwa hasil pengelompokan kasus pasien yang sembuh dan meninggal dari COVID-19 kabupaten-kota di Provinsi Jawa Timur terbagi menjadi 4 *cluster*. Dari keseluruhan 38 kabupaten-kota Provinsi

Jawa Timur, terbagi pada *cluster* 0 terdiri dari 1 kota di Jawa Timur, *cluster* 1 terdiri dari 9 kabupaten dan kota di Jawa Timur dan *cluster* 2 terdiri dari 13 kabupaten dan kota di Jawa Timur dan *cluster* 3 terdiri dari 15 kabupaten dan kota di Jawa Timur. Kemudian langkah selanjutnya yakni menentukan ciri/karakteristik dari setiap kelompok yang terbentuk.

4.7 Ciri Kelompok

Untuk mengetahui ciri/karakteristik dari masing-masing *cluster*, penelitian ini menggunakan *scatter plot* dengan bantuan *software* minitab. *Scatter plot* merupakan sebuah tipe grafik yang digunakan untuk menggambarkan suatu data dengan menggunakan koordinat kartesius. Data-data yang ditampilkan pada *scatter plot* baik digunakan untuk melihat relasi yang terbentuk antar 2 variabel dengan melihat keberadaan titik variabel pada koordinat X dan Y. Adapun untuk menggambarkan kondisi suatu wilayah, Badan Nasional Penanggulangan Bencana (BNPB) membagi kategori wilayah kedalam 4 warna dalam peta zona resiko, yakni: hijau, kuning, orange, dan merah. Adapun penjelasan dari keempat warna tersebut adalah sebagai berikut:

a Zona hijau

Wilayah tanpa kasus dikonfirmasi atau dengan beberapa kasus infeksi yang datang dari negara lain

b Zona kuning

Wilayah dengan beberapa kasus penularan lokal, tanpa adanya penularan dalam bentuk komunitas.

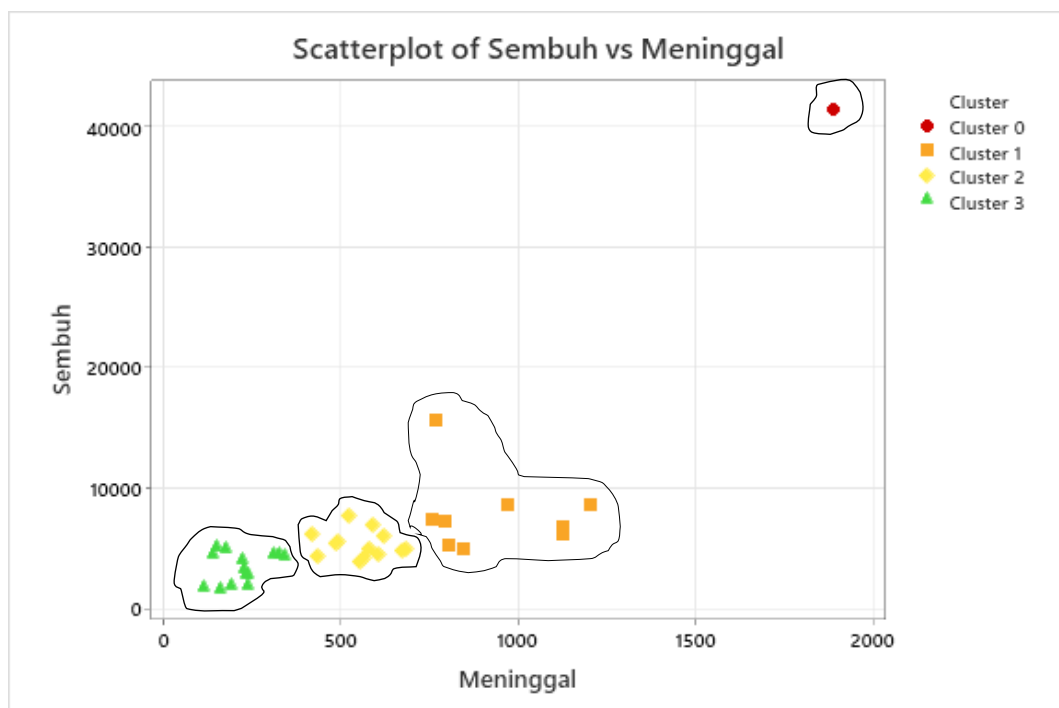
c Zona orange

Wilayah berdekatan dengan zona merah atau dengan kluster penyebaran kecil.

d Zona Merah

Wilayah dengan banyak kasus dikonfirmasi.

Dengan menggunakan warna identifikasi wilayah menurut Badan Nasional Penanggulangan Bencana (BNPB) pengelompokan di visualisasikan menggunakan *scatterplot* sebagai berikut:



Gambar 4.6 Scatterplot Meninggal Vs Sembuh

(Sumber: Output Minitab)

Berdasarkan output pada Gambar 4.6 menunjukkan perbedaan antara 4 *cluster* yang terbentuk oleh *k-means cluster* masing-masing *cluster* tersebut memiliki karakteristik (ciri kelompok) sebagai berikut:

1. Titik berwarna merah terdiri dari anggota kelompok dari *cluster* 0 yakni: Kota Surabaya. Berdasarkan *scatterplot* yang terbentuk, dapat diketahui

bahwa *cluster* 1 merupakan kelompok yang memiliki ciri/karakteristik jumlah pasien meninggal dan sembuh paling banyak diantara *cluster* 1, *cluster* 2 dan *cluster* 3 atau dapat dikategorikan *cluster* 0 ini memiliki jumlah pasien sembuh dan meninggal akibat COVID-19 “sangat tinggi” di Provinsi Jawa Timur.

2. Titik berwarna orange terdiri dari anggota kelompok dari *cluster* 1. Berdasarkan *scatterplot* yang terbentuk, dapat diketahui bahwa kelompok ini memiliki ciri/karakteristik jumlah pasien tinggi, artinya kasus sembuh dan meninggal akibat COVID-19 lebih kecil dari *cluster* 0, namun lebih besar dari *cluster* 2 dan *cluster* 3, dengan rata-rata kasus sembuh 7882 pasien dan kasus meninggal 931 pasien. Sehingga *cluster* 1 ini dikategorikan memiliki jumlah pasien sembuh dan meninggal akibat COVID-19 “tinggi” di Provinsi Jawa Timur.
3. Titik berwarna kuning terdiri dari anggota kelompok dari *cluster* 2. Berdasarkan *scatterplot* yang terbentuk, dapat diketahui bahwa kelompok ini memiliki ciri/karakteristik jumlah pasien rendah, artinya kasus sembuh dan meninggal akibat COVID-19 lebih sedikit dari *cluster* 0 dan *cluster* 1, namun lebih banyak dari *cluster* 3 dengan rata-rata kasus sembuh 5286 pasien dan kasus meninggal 555 pasien. Sehingga *cluster* 2 ini dikategorikan memiliki jumlah pasien sembuh dan meninggal akibat COVID-19 “cukup/sedang” di Provinsi Jawa Timur.
4. Titik berwarna hijau terdiri dari anggota kelompok dari *cluster* 3. Berdasarkan *scatterplot* yang terbentuk, dapat diketahui bahwa *cluster* 3 merupakan kelompok yang memiliki ciri/karakteristik jumlah pasien

meninggal dan sembuh paling sedikit dari *cluster 1*, *cluster 2* dan *cluster 3* dengan rata-rata kasus sembuh 3481 pasien dan kasus meninggal 216 pasien. Sehingga dapat dikategorikan *cluster 3* ini memiliki jumlah pasien sembuh dan meninggal akibat COVID-19 “rendah” di Provinsi Jawa Timur.

4.8 Kajian Keislaman dengan Hasil Penelitian

Berdasarkan hasil yang diperoleh dari penelitian ini yakni pengelempokan menjadi 4 wilayah kabupaten-kota di Jawa Timur yang memiliki karakteristik hampir sama merupakan salah satu upaya untuk mengantisipasi dan meminimalisir penyebaran COVID-19. Mengingat besar dan luasnya Provinsi Jawa Timur, pengelompokan wilayah juga dapat membantu untuk memberikan wawasan kepada masyarakat serta mempermudah penanganan pasien terkonfirmasi COVID-19, seperti yang diriwayatkan dalam hadits (Supriatna, 2020):

عَنْ عَبْدِ اللَّهِ بْنِ عَمْرٍو بْنِ رَبِيعَةَ أَنَّ عُمَرَ خَرَجَ إِلَى الشَّامِ فَلَمَّا جَاءَ سَرَّحَ بَلَعَهُ أَنَّ الْوَبَاءَ قَدْ وَقَعَ بِالشَّامِ فَأَخْبَرَهُ عَبْدُ الرَّحْمَنِ بْنُ عَوْفٍ أَنَّ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ إِذَا سَمِعْتُمْ بِهِ بِأَرْضٍ فَلَا تَقْدَمُوا عَلَيْهِ وَإِذَا وَقَعَ بِأَرْضٍ وَأَنْتُمْ بِهَا فَلَا تَخْرُجُوا فِرَارًا مِنْهُ فَرَجَعَ عُمَرُ بْنُ الْخَطَّابِ مِنْ سَرَّحَ

Artinya:

“Dari Abdullah bin Amir bin Rabi’ah, Umar bin Khattab RA menempuh perjalanan menuju Syam. Ketika sampai di Sargh, Umar mendapat kabar bahwa wabah sedang menimpa wilayah Syam. Abdurrahman bin Auf mengatakan kepada Umar bahwa Rasulullah SAW pernah bersabda, ‘Bila kamu mendengar wabah di suatu daerah, maka kalian jangan memasukinya. Tetapi jika wabah terjadi wabah di daerah kamu berada, maka jangan tinggalkan tempat itu.’ Lalu Umar bin Khattab berbalik arah meninggalkan Sargh” (HR Bukhari dan Muslim).

Dalam hadits tersebut disebutkan apabila terdapat suatu tempat yang terdampak suatu wabah (COVID-19), hendaklah tetap berada ditempat tersebut dan jangan pergi ke wilayah lain. Langkah yang dilakukan menurut hadits diatas adalah

dengan menghindar/menjaga jarak terhadap orang yang sedang menderita virus tersebut. Anjuran untuk menjaga jarak tersebut dikarenakan wabah COVID-19 termasuk dalam kategori virus yang menular dengan sangat cepat, sehingga masyarakat diharuskan untuk menjalankan kebijakan baru seperti *Social Distancing* dan *Work From Home* (WFH) yakni pembatasan dalam kegiatan sehari-hari dengan melakukan aktivitas di rumah masing-masing, baik berupa aktivitas kerja, sekolah, dan lain-lain. Kebijakan pembatasan tersebut dilakukan dengan aturan yang berbeda pada tiap-tiap *cluster* dengan melihat kondisi masyarakat pada wilayah pada titik merah, titik orange, titik kuning dan titik hijau. Selain itu Imam Bukhari meriwayatkan dari sahabat Abu Hurairah:

حَدَّثَنَا مُحَمَّدُ بْنُ بَشَّارٍ حَدَّثَنَا عَبْدُ الرَّحْمَنِ بْنُ مَهْدِيٍّ حَدَّثَنَا سُفْيَانُ عَنْ عَطَاءِ بْنِ السَّائِبِ عَنْ أَبِي عَبْدِ الرَّحْمَنِ عَنْ عَبْدِ اللَّهِ عَنْ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ مَا أَنْزَلَ اللَّهُ دَاءً إِلَّا أَنْزَلَ لَهُ دَوَاءً

Artinya:

“Telah menceritakan kepada kami Muhammad bin Basysyar telah menceritakan kepada kami Abdurrahman bin Mahdi telah menceritakan kepada kami Sufyan dari ‘Atha bin As Sa’ib dari Abdurrahman dari Abdullah dari Nabi shallallahu ‘alaihi wasallam, beliau bersabda: Sesungguhnya Allah tidak menurunkan suatu penyakit kecuali diturunkan juga baginya obat”

Dari hadits diatas dapat disimpulkan bahwa ketika Allah memberikan penyakit, maka Dia juga membuat obatnya. Hal ini juga berlaku pada orang yang terpapar virus COVID-19. Tentunya untuk menemukan obat tersebut harus dicari dengan usaha dan penuh kesabaran sebab proses penyembuhan membutuhkan waktu yang tidak cepat. Ikhtiar untuk mendapatkan obat tersebut harus diimbangi dengan usaha merubah pola hidup yang sehat, cuci tangan, menggunakan masker, serta membatasi interaksi kepada orang lain dengan cara isolasi selama kurun waktu kurang lebih 14 hari.

BAB V PENUTUP

5.1 Kesimpulan

Sesuai dengan penelitian yang telah dilakukan oleh peneliti pada data kasus COVID-19 di Provinsi Jawa Timur dengan menggunakan analisis *k-means cluster* metode *elbow*, didapatkan kesimpulan sebagai berikut:

1. Pengelompokan kasus sembuh dan kasus meninggal akibat COVID-19 dengan menggunakan analisis *k-means cluster* metode *elbow* menghasilkan 4 *cluster*, dengan perincian: *cluster* 0 terdiri dari 1 kota, *cluster* 1 terdiri dari 9 kabupaten-kota, *cluster* 2 terdiri dari 13 kabupaten-kota dan *cluster* 3 terdiri dari 15 kabupaten-kota di Provinsi Jawa Timur
2. Masing-masing *cluster* memiliki ciri kelompok sebagai berikut:
 - a. *Cluster* 0 (Titik Merah) memiliki ciri jumlah pasien yang sembuh akibat COVID-19 (X1) dan meninggal akibat COVID-19 (X2) pada tingkat sangat tinggi.
 - b. *Cluster* 1 (Titik Orange) memiliki ciri jumlah pasien yang sembuh akibat COVID-19 (X1) dan meninggal akibat COVID-19 (X2) pada tingkat tinggi.
 - c. *Cluster* 2 (Titik Kuning) memiliki ciri pasien yang sembuh akibat COVID-19 (X1) dan meninggal akibat COVID-19 (X2) pada tingkat cukup/sedang.
 - d. *Cluster* 3 (Titik Hijau) memiliki ciri jumlah pasien yang sembuh akibat COVID-19 (X1) dan meninggal akibat COVID-19 (X2) pada tingkat rendah.

5.2 Saran

Pada penelitian ini peneliti hanya mengkaji teknik pengelompokan *non-hierarki k-means cluster* dengan pemilihan *k optimal* menggunakan metode *elbow* yang diaplikasikan pada data kasus COVID-19. Oleh sebab itu, bagi pembaca yang akan melakukan penelitian *clustering* dapat menggunakan metode atau variabel yang lain.

DAFTAR PUSTAKA

- Al-Quran Terjemahan Departemen Agama RI. (2015). Bandung: CV. Darus Sunnah.
- Berry, M. (2016). *Soft Computing in Data Science*. Singapore: Springer Nature.
- Burhan, E., & dkk. (2020). *Pneumonia COVID-19 Diagnosa & Penatalaksanaan di Indonesia*. Jakarta: Perhimpunan Dokter Paru Indonesia (PDPI).
- Ernawati, Merliana, N. P., & Santoso, A. (2015). Analisa Penentuan Cluster Terbaik Pada Metode K-Means Clustering. *Prosiding Seminar Nasional Multi Disiplin Ilmu & Call For Paper UNISBANK (SENDI_U)*, 1-7.
- Ghoffar, M. A., & Al-Atsari, A. I. (2004). *Tafsir Ibnu Katsir Jilid 7* (7 ed.). Bogor: Pustaka Imam Asy-Syafi'i.
- Hair, J., Black, W., & Babin, B. (2009). *Multivariate Data Analysis Seventh Edition*. Upper Saddle River: Prentice Hall.
- Han, J., & Kamber, M. (2001). *Data Mining Concept And Techniques*. USA: Academic Press.
- Harani, H. N. (2020). *Segmentasi Pelanggan Menggunakan Python*. Bandung: Kreatif Industri Nusantara.
- Johnson, R., & Winchern, D. (2007). *Applied Multivariate Statistical Analysis 6th Edition*. London: Prentice-Hall.
- Karyono, Indriyani, D., & Rohadin, R. (2020). Penanganan Pencegahan Pandemi Wabah Virus Corona (Covid-19) Kabupaten Indramayu. *Jurnal Kolaborasi Resolusi Konflik, II*(2), 164-173.
- Pemprov Jatim. (2021). *Jatim Tanggap Covid-19*. Pemerintah Provinsi Jawa Timur. Dipetik Maret 11, 2021, dari <http://infocovid19.jatimprov.go.id/>
- Prakasa, R., & Kurnianingtyas, A. (2022). *SPSS untuk Analisis & Perencanaan Kota yang Lebih Baik*. Yogyakarta: CV Andi Offset.
- Rahman, A. T., Wiranto, & Anggrainingsih, R. (2017). Coal Trade Data Clustering Using K-Means. *Jurnal Ilmiah Teknologi dan Informasi, VI*(1), 24-31.
- Relia Soft Corporation. (2018). *Multiple Linear Regression Analysis*. Dipetik Agustus 15, 2021, dari http://reliawiki.org/index.php/Multiple_Linear_Regression_Analysis

- Setiawan, A. R. (2020). Lembar Kegiatan Literasi Saintifik untuk Pembelajaran Jarak Jauh Topik Penyakit Coronavirus 2019 (COVID-19). *Jurnal Edukatif, II*, 28-37.
- Shihab, M. Q. (2002). *Tafsir Al Misbah : pesan,kesan dan keserasian Al-Qur'an* (Vol. 1). Jakarta: Lentera Hati.
- Soetcipto, d. (2020). *Memahami Perilaku Covid-19 di Jawa Timur*. Surabaya: Airlangga University Press.
- Supriatna, E. (2020). Wabah Corona Virus Disease Covid 19 dalam Pandangan Islam. *Jurnal Sosial dan Budaya Syar'i*, 7 (6), 555-564.
- Wahyudi, M. (2020). *Data Mining: Penerapan Algoritma K-Means Clustering dan K-Medoids Clustering*. Yayasan Kita Menulis.
- WHO. (2020). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*. World Health Organization. Dipetik Maret 11, 2021, dari <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
- Yudiatmaja, F. (2013). *Analisis Regresi dengan Menggunakan Aplikasi Komputer Statistik*. Jakarta: PT Gramedia Pustaka Utama.
- Yunus, N. R., & Rezki, A. (2020). Kebijakan Pemberlakuan Lockdown Sebagai Antisipasi Penyebaran Corona Virus Covid-19. *Jurnal Sosial & Budaya Syar-i*, 7, 227-238.
- Zurraedah. (2018). *Evaluasi Klasifikasi Kategori Pembayaran SPP di Universitas Islam Negeri Alauddin Makassar (UINAM) Tahun Akademik 2017/2018 Menggunakan K-Means CLustering*. Makassar: Repository UIN Alauddin Makassar.

LAMPIRAN

Lampiran 1. Data Persebaran COVID-19 per 31 Juli 2021

(Sumber : <http://infocovid19.jatimprov.go.id>)

Kabupaten-Kota	Kasus COVID-19		Kabupaten-Kota	Kasus COVID-19	
	Sembuh	Meninggal		Sembuh	Meninggal
PACITAN	5095	146	MAGETAN	5992	619
PONOROGO	5201	803	NGAWI	4957	580
TRENGGALEK	4739	674	BOJONEGORO	4289	432
TULUNGAGUNG	4484	138	TUBAN	5013	843
BLITAR	6266	1126	LAMONGAN	4393	310
KEDIRI	7410	754	GRESIK	6876	587
MALANG	6134	419	BANGKALAN	4410	602
LUMAJANG	4101	561	SAMPANG	1674	110
JEMBER	8664	968	PAMEKASAN	1488	159
BANYUWANGI	8589	1201	SUMENEP	4069	220
BONDOWOSO	3876	552	KOTA KEDIRI	1883	234
SITUBONDO	4853	684	KOTA BLITAR	4880	171
PROBOLINGGO	4451	325	KOTA MALANG	7240	791
PASURUAN	5528	492	KOTA PROBOLINGGO	2716	236
SIDOARJO	15710	766	KOTA PASURUAN	2747	229
MOJOKERTO	4854	175	KOTA MOJOKERTO	3270	224
JOMBANG	6842	1126	KOTA MADIUN	4337	338
NGANJUK	7642	519	KOTA SURABAYA	41448	1888
MADIUN	5318	483	KOTA BATU	1873	188

Lampiran 2. Matriks H
(Matriks dengan ordo 38x38)

No	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0.027	0.027	0.027	0.027	0.026	0.025	0.026	0.028	0.025	0.025	0.028	0.027	0.027
2	0.027	0.027	0.027	0.027	0.026	0.026	0.026	0.028	0.025	0.025	0.028	0.027	0.027
3	0.027	0.027	0.028	0.028	0.026	0.025	0.026	0.028	0.024	0.024	0.028	0.028	0.028
4	0.027	0.027	0.028	0.028	0.026	0.025	0.026	0.029	0.024	0.024	0.029	0.028	0.028
5	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.027	0.027	0.026	0.026	0.026
6	0.025	0.026	0.025	0.025	0.026	0.027	0.026	0.025	0.028	0.028	0.024	0.025	0.025
7	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026
8	0.028	0.028	0.028	0.029	0.026	0.025	0.026	0.029	0.023	0.023	0.029	0.028	0.029
9	0.025	0.025	0.024	0.024	0.027	0.028	0.026	0.023	0.031	0.03	0.023	0.024	0.023
10	0.025	0.025	0.024	0.024	0.027	0.028	0.026	0.023	0.03	0.03	0.023	0.024	0.024
11	0.028	0.028	0.028	0.029	0.026	0.024	0.026	0.029	0.023	0.023	0.03	0.028	0.029
12	0.027	0.027	0.028	0.028	0.026	0.025	0.026	0.028	0.024	0.024	0.028	0.027	0.028
13	0.027	0.027	0.028	0.028	0.026	0.025	0.026	0.029	0.023	0.024	0.029	0.028	0.028
14	0.027	0.027	0.027	0.027	0.026	0.026	0.026	0.027	0.025	0.025	0.027	0.027	0.027
15	0.02	0.02	0.017	0.016	0.027	0.034	0.026	0.013	0.042	0.042	0.012	0.018	0.016
16	0.027	0.027	0.028	0.028	0.026	0.025	0.026	0.028	0.024	0.024	0.028	0.027	0.028
17	0.026	0.026	0.026	0.026	0.026	0.027	0.026	0.025	0.027	0.027	0.025	0.026	0.026
18	0.025	0.025	0.025	0.025	0.026	0.028	0.026	0.024	0.029	0.029	0.024	0.025	0.025
19	0.027	0.027	0.027	0.027	0.026	0.026	0.026	0.027	0.025	0.025	0.028	0.027	0.027
20	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.027	0.026	0.026	0.027	0.026	0.026
21	0.027	0.027	0.027	0.028	0.026	0.025	0.026	0.028	0.024	0.024	0.028	0.027	0.028
22	0.028	0.027	0.028	0.028	0.026	0.025	0.026	0.029	0.023	0.023	0.029	0.028	0.028
23	0.027	0.027	0.027	0.028	0.026	0.025	0.026	0.028	0.024	0.024	0.028	0.027	0.028
24	0.028	0.027	0.028	0.028	0.026	0.025	0.026	0.029	0.023	0.023	0.029	0.028	0.028
25	0.026	0.026	0.026	0.026	0.026	0.027	0.026	0.025	0.028	0.028	0.025	0.026	0.025
26	0.028	0.027	0.028	0.028	0.026	0.025	0.026	0.029	0.023	0.024	0.029	0.028	0.028
27	0.029	0.029	0.03	0.031	0.026	0.023	0.026	0.032	0.019	0.019	0.033	0.03	0.031
28	0.03	0.029	0.031	0.031	0.026	0.022	0.026	0.033	0.019	0.019	0.033	0.03	0.032
29	0.028	0.028	0.028	0.029	0.026	0.025	0.026	0.029	0.023	0.023	0.029	0.028	0.029
30	0.029	0.029	0.03	0.031	0.026	0.023	0.026	0.032	0.019	0.019	0.033	0.03	0.031
31	0.027	0.027	0.027	0.028	0.026	0.025	0.026	0.028	0.024	0.024	0.028	0.027	0.028
32	0.026	0.026	0.025	0.025	0.026	0.027	0.026	0.025	0.028	0.028	0.025	0.025	0.025
33	0.029	0.028	0.029	0.03	0.026	0.023	0.026	0.031	0.021	0.021	0.031	0.029	0.03
34	0.029	0.028	0.029	0.03	0.026	0.023	0.026	0.031	0.021	0.021	0.031	0.029	0.03
35	0.028	0.028	0.029	0.029	0.026	0.024	0.026	0.03	0.022	0.022	0.031	0.029	0.03
36	0.028	0.027	0.028	0.028	0.026	0.025	0.026	0.029	0.023	0.023	0.029	0.028	0.028
37	0.002	0.004	-0.006	-0.012	0.029	0.056	0.026	-0.02	0.085	0.084	-0.027	-0	-0.01
38	0.029	0.029	0.03	0.031	0.026	0.023	0.026	0.032	0.019	0.019	0.033	0.03	0.031

Lampiran 2. (lanjutan Matriks H)

No	14	15	16	17	18	19	20	21	22	23	24	25
1	0.027	0.02	0.027	0.026	0.025	0.027	0.026	0.027	0.028	0.027	0.028	0.026
2	0.027	0.02	0.027	0.026	0.025	0.027	0.026	0.027	0.027	0.027	0.027	0.026
3	0.027	0.017	0.028	0.026	0.025	0.027	0.026	0.027	0.028	0.027	0.028	0.026
4	0.027	0.016	0.028	0.026	0.025	0.027	0.026	0.028	0.028	0.028	0.028	0.026
5	0.026	0.027	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026
6	0.026	0.034	0.025	0.027	0.028	0.026	0.026	0.025	0.025	0.025	0.025	0.027
7	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026
8	0.027	0.013	0.028	0.025	0.024	0.027	0.027	0.028	0.029	0.028	0.029	0.025
9	0.025	0.042	0.024	0.027	0.029	0.025	0.026	0.024	0.023	0.024	0.023	0.028
10	0.025	0.042	0.024	0.027	0.029	0.025	0.026	0.024	0.023	0.024	0.023	0.028
11	0.027	0.012	0.028	0.025	0.024	0.028	0.027	0.028	0.029	0.028	0.029	0.025
12	0.027	0.018	0.027	0.026	0.025	0.027	0.026	0.027	0.028	0.027	0.028	0.026
13	0.027	0.016	0.028	0.026	0.025	0.027	0.026	0.028	0.028	0.028	0.028	0.025
14	0.027	0.022	0.027	0.026	0.026	0.027	0.026	0.027	0.027	0.027	0.027	0.026
15	0.022	0.087	0.018	0.031	0.036	0.021	0.025	0.019	0.015	0.019	0.015	0.031
16	0.027	0.018	0.027	0.026	0.025	0.027	0.026	0.027	0.028	0.027	0.028	0.026
17	0.026	0.031	0.026	0.027	0.027	0.026	0.026	0.026	0.025	0.026	0.026	0.027
18	0.026	0.036	0.025	0.027	0.028	0.025	0.026	0.025	0.024	0.025	0.025	0.027
19	0.027	0.021	0.027	0.026	0.025	0.027	0.026	0.027	0.027	0.027	0.027	0.026
20	0.026	0.025	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026
21	0.027	0.019	0.027	0.026	0.025	0.027	0.026	0.027	0.028	0.027	0.028	0.026
22	0.027	0.015	0.028	0.025	0.024	0.027	0.026	0.028	0.029	0.028	0.028	0.025
23	0.027	0.019	0.027	0.026	0.025	0.027	0.026	0.027	0.028	0.027	0.028	0.026
24	0.027	0.015	0.028	0.026	0.025	0.027	0.026	0.028	0.028	0.028	0.028	0.025
25	0.026	0.031	0.026	0.027	0.027	0.026	0.026	0.026	0.025	0.026	0.025	0.027
26	0.027	0.015	0.028	0.026	0.025	0.027	0.026	0.028	0.028	0.028	0.028	0.025
27	0.028	-0.002	0.03	0.024	0.022	0.029	0.027	0.03	0.032	0.03	0.031	0.024
28	0.028	-0.003	0.03	0.024	0.022	0.029	0.027	0.03	0.032	0.03	0.032	0.024
29	0.027	0.013	0.028	0.025	0.024	0.027	0.027	0.028	0.029	0.028	0.029	0.025
30	0.028	-7E-04	0.03	0.024	0.022	0.029	0.027	0.03	0.032	0.029	0.031	0.024
31	0.027	0.018	0.027	0.026	0.025	0.027	0.026	0.027	0.028	0.027	0.028	0.026
32	0.026	0.033	0.025	0.027	0.027	0.026	0.026	0.025	0.025	0.025	0.025	0.027
33	0.028	0.005	0.029	0.025	0.023	0.028	0.027	0.029	0.031	0.029	0.03	0.025
34	0.028	0.005	0.029	0.025	0.023	0.028	0.027	0.029	0.03	0.029	0.03	0.025
35	0.027	0.008	0.029	0.025	0.023	0.028	0.027	0.029	0.03	0.028	0.03	0.025
36	0.027	0.015	0.028	0.025	0.025	0.027	0.026	0.028	0.029	0.028	0.028	0.025
37	0.012	0.25	-0.004	0.043	0.061	0.007	0.023	-0.001	-0.02	-0	-0.01	0.044

Lampiran 2. (lanjutan Matriks H)

No	26	27	28	29	30	31	32	33	34	35	36	37	38
1	0.028	0.029	0.03	0.028	0.029	0.027	0.026	0.029	0.029	0.028	0.028	0.002	0.029
2	0.027	0.029	0.029	0.028	0.029	0.027	0.026	0.028	0.028	0.028	0.027	0.004	0.029
3	0.028	0.03	0.031	0.028	0.03	0.027	0.025	0.029	0.029	0.029	0.028	-0.01	0.03
4	0.028	0.031	0.031	0.029	0.031	0.028	0.025	0.03	0.03	0.029	0.028	-0.01	0.031
5	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.029	0.026
6	0.025	0.023	0.022	0.025	0.023	0.025	0.027	0.023	0.023	0.024	0.025	0.056	0.023
7	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026	0.026
8	0.029	0.032	0.033	0.029	0.032	0.028	0.025	0.031	0.031	0.03	0.029	-0.02	0.032
9	0.023	0.019	0.019	0.023	0.019	0.024	0.028	0.021	0.021	0.022	0.023	0.085	0.019
10	0.024	0.019	0.019	0.023	0.019	0.024	0.028	0.021	0.021	0.022	0.023	0.084	0.019
11	0.029	0.033	0.033	0.029	0.033	0.028	0.025	0.031	0.031	0.031	0.029	-0.03	0.033
12	0.028	0.03	0.03	0.028	0.03	0.027	0.025	0.029	0.029	0.029	0.028	-0	0.03
13	0.028	0.031	0.032	0.029	0.031	0.028	0.025	0.03	0.03	0.03	0.028	-0.01	0.031
14	0.027	0.028	0.028	0.027	0.028	0.027	0.026	0.028	0.028	0.027	0.027	0.012	0.028
15	0.015	-0	-0	0.013	-0	0.018	0.033	0.005	0.005	0.008	0.015	0.25	-0
16	0.028	0.03	0.03	0.028	0.03	0.027	0.025	0.029	0.029	0.029	0.028	-0	0.03
17	0.026	0.024	0.024	0.025	0.024	0.026	0.027	0.025	0.025	0.025	0.025	0.043	0.024
18	0.025	0.022	0.022	0.024	0.022	0.025	0.027	0.023	0.023	0.023	0.025	0.061	0.022
19	0.027	0.029	0.029	0.027	0.029	0.027	0.026	0.028	0.028	0.028	0.027	0.007	0.029
20	0.026	0.027	0.027	0.027	0.027	0.026	0.026	0.027	0.027	0.027	0.026	0.023	0.027
21	0.028	0.03	0.03	0.028	0.03	0.027	0.025	0.029	0.029	0.029	0.028	-0	0.03
22	0.028	0.032	0.032	0.029	0.032	0.028	0.025	0.031	0.03	0.03	0.029	-0.02	0.032
23	0.028	0.03	0.03	0.028	0.029	0.027	0.025	0.029	0.029	0.028	0.028	-0	0.029
24	0.028	0.031	0.032	0.029	0.031	0.028	0.025	0.03	0.03	0.03	0.028	-0.01	0.031
25	0.025	0.024	0.024	0.025	0.024	0.026	0.027	0.025	0.025	0.025	0.025	0.044	0.024
26	0.028	0.031	0.032	0.029	0.031	0.028	0.025	0.03	0.03	0.03	0.028	-0.01	0.031
27	0.031	0.04	0.04	0.032	0.039	0.03	0.023	0.036	0.036	0.035	0.032	-0.08	0.039
28	0.032	0.04	0.041	0.033	0.039	0.03	0.023	0.037	0.037	0.035	0.032	-0.08	0.039
29	0.029	0.032	0.033	0.029	0.032	0.028	0.025	0.031	0.031	0.03	0.029	-0.02	0.032
30	0.031	0.039	0.039	0.032	0.038	0.03	0.023	0.036	0.036	0.034	0.031	-0.07	0.038
31	0.028	0.03	0.03	0.028	0.03	0.027	0.025	0.029	0.029	0.029	0.028	-0	0.03
32	0.025	0.023	0.023	0.025	0.023	0.025	0.027	0.024	0.024	0.024	0.025	0.052	0.023
33	0.03	0.036	0.037	0.031	0.036	0.029	0.024	0.034	0.034	0.033	0.03	-0.05	0.036
34	0.03	0.036	0.037	0.031	0.036	0.029	0.024	0.034	0.034	0.033	0.03	-0.05	0.036
35	0.03	0.035	0.035	0.03	0.034	0.029	0.024	0.033	0.033	0.032	0.03	-0.04	0.034
36	0.028	0.032	0.032	0.029	0.031	0.028	0.025	0.03	0.03	0.03	0.028	-0.02	0.031
37	-0.01	-0.08	-0.08	-0.022	-0.07	-0.003	0.052	-0.054	-0.05	-0.04	-0.02	0.852	-0.07
38	0.031	0.039	0.039	0.032	0.038	0.03	0.023	0.036	0.036	0.034	0.031	-0.07	0.038

Lampiran 5. Data Setelah Dinormalisasi

Kabupaten-Kota	Kasus COVID-19		Kabupaten-Kota	Kasus COVID-19	
	Sembuh	Meninggal		Sembuh	Meninggal
PACITAN	0.090265265	0.020247469	MAGETAN	0.112712713	0.286276715
PONOROGO	0.092917918	0.38976378	NGAWI	0.086811812	0.264341957
TRENGGALEK	0.081356356	0.317210349	BOJONEGORO	0.070095095	0.181102362
TULUNGAGUNG	0.074974975	0.015748031	TUBAN	0.088213213	0.412260967
BLITAR	0.11956957	0.571428571	LAMONGAN	0.072697698	0.112485939
KEDIRI	0.148198198	0.362204724	GRESIK	0.134834835	0.268278965
MALANG	0.116266266	0.173790776	BANGKALAN	0.073123123	0.276715411
LUMAJANG	0.06539039	0.253655793	SAMPANG	0.004654655	0
JEMBER	0.17957958	0.482564679	PAMEKASAN	0	0.027559055
BANYUWANGI	0.177702703	0.613610799	SUMENEP	0.06458959	0.061867267
BONDOWOSO	0.05975976	0.248593926	KOTA KEDIRI	0.009884885	0.069741282
SITUBONDO	0.084209209	0.322834646	KOTA BLITAR	0.084884885	0.034308211
PROBOLINGGO	0.074149149	0.120922385	KOTA MALANG	0.143943944	0.383014623
PASURUAN	0.101101101	0.214848144	KOTA PROBOLINGGO	0.030730731	0.070866142
SIDOARJO	0.355905906	0.368953881	KOTA PASURUAN	0.031506507	0.066929134
MOJOKERTO	0.084234234	0.03655793	KOTA MOJOKERTO	0.044594595	0.064116985
JOMBANG	0.133983984	0.571428571	KOTA MADIUN	0.071296296	0.128233971
NGANJUK	0.154004004	0.230033746	KOTA SURABAYA	1	1
MADIUN	0.095845846	0.209786277	KOTA BATU	0.009634635	0.043869516

Lampiran 6. Hasil Perhitungan *Euclidian Distance* pada Iterasi-1

No	C0	C1	C2	C3	Jarak Minimum	Cluster yang diikuti
1	1.336986	0.156315	0.034592	0.092508	0.034592	2
2	1.093246	0.217611	0.354759	0.400686	0.217611	1
3	1.144599	0.148423	0.281482	0.327477	0.148423	1
4	1.350712	0.164214	0.025779	0.076611	0.025779	2
5	0.979199	0.397638	0.538204	0.583804	0.397638	1
6	1.064119	0.190577	0.337442	0.39135	0.190577	1
7	1.209796	0.003303	0.148411	0.209096	0.003303	1
8	1.196045	0.096508	0.217171	0.261949	0.096508	1
9	0.969963	0.314551	0.461821	0.514896	0.314551	1
10	0.908553	0.443645	0.588983	0.638824	0.443645	1
11	1.203604	0.095774	0.212036	0.255676	0.095774	1
12	1.138958	0.153181	0.287319	0.333637	0.153181	1
13	1.276706	0.0697	0.085583	0.141846	0.0697	1
14	1.193517	0.04502	0.183021	0.237447	0.04502	1
15	0.901707	0.306502	0.445185	0.512636	0.306502	1
16	1.329228	0.141709	0.024474	0.091825	0.024474	2
17	0.966259	0.397899	0.539996	0.586926	0.397899	1
18	1.143922	0.065947	0.215209	0.276826	0.065947	1
19	1.200805	0.04311	0.176947	0.230644	0.04311	1
20	1.138718	0.112695	0.255271	0.307666	0.112695	1
21	1.172649	0.096294	0.229385	0.278232	0.096294	1
22	1.239079	0.050012	0.144913	0.194194	0.050012	1
23	1.084801	0.240523	0.376779	0.421593	0.240523	1
24	1.283577	0.07717	0.077022	0.133933	0.077022	2
25	1.133105	0.095713	0.243579	0.300257	0.095713	1
26	1.175688	0.112919	0.240529	0.286214	0.112919	1
27	1.410926	0.208347	0.066129	0.004655	0.004655	3
28	1.394862	0.188893	0.060434	0.027559	0.027559	3
29	1.324796	0.124698	0.025766	0.089439	0.025766	2
30	1.358569	0.151185	0.059905	0.070438	0.059905	2
31	1.330412	0.14373	0.025226	0.091556	0.025226	2
32	1.055226	0.210639	0.356538	0.40917	0.210639	1
33	1.342674	0.135963	0.044941	0.077242	0.044941	2
34	1.344842	0.138472	0.041481	0.073974	0.041481	2
35	1.337414	0.132852	0.031456	0.0781	0.031456	2
36	1.273761	0.066376	0.092399	0.146721	0.066376	1
37	0	1.207385	1.346207	1.414214	0	0
38	0.376593	0.070092	0.050656	0.044905	0.044905	3

Lampiran 7. Hasil Perhitungan *Euclidian Distance* pada Iterasi-2

No	C0	C0	C2	C3	Jarak Minimum	Cluster yang diikuti
0	0.336986	0.299823	0.047069	0.085576	0.04706947	2
2	0.093246	0.075088	0.336209	0.376422	0.075088273	1
3	0.044599	0.03622	0.26289	0.303234	0.036220203	1
4	0.350702	0.306054	0.042706	0.070673	0.042705655	2
5	0.979099	0.252609	0.509703	0.559524	0.252609387	1
6	0.064009	0.053008	0.309663	0.367539	0.053007747	1
7	0.209796	0.045042	0.030687	0.086889	0.030686673	2
8	0.096045	0.083468	0.098477	0.237708	0.083468303	1
9	0.969963	0.075096	0.44400	0.490935	0.075096462	1
10	0.908553	0.30086	0.570837	0.604633	0.300860079	1
11	0.203604	0.090947	0.093309	0.230404	0.090946900	1
12	0.038958	0.033570	0.268748	0.309399	0.033570277	1
13	0.276706	0.202606	0.067398	0.009354	0.067398024	2
14	0.093507	0.005270	0.065064	0.203955	0.00527072	1
15	0.900707	0.243579	0.432005	0.492368	0.243579059	1
16	0.329228	0.284228	0.030557	0.080487	0.030556699	2
17	0.966259	0.253036	0.520584	0.562659	0.253035880	1
18	0.043922	0.095989	0.098980	0.254560	0.095989008	1
19	0.200805	0.000078	0.05887	0.207083	0.000078273	1
20	0.038708	0.032907	0.23709	0.283799	0.032906685	1
20	0.072649	0.062553	0.200909	0.254040	0.062553225	1
22	0.239079	0.045668	0.026308	0.070320	0.026308270	2
23	0.084800	0.097928	0.358080	0.397304	0.097928093	1
24	0.283577	0.200058	0.058855	0.000708	0.058854689	2
25	0.033005	0.053425	0.226045	0.276909	0.053425023	1
26	0.075688	0.060207	0.220889	0.260982	0.060207037	1
27	0.400926	0.338222	0.07740	0.02380	0.023809770	3
28	0.394862	0.304090	0.065043	0.006062	0.006060854	3
29	0.324796	0.262359	0.008740	0.070906	0.008740820	2
30	0.358569	0.270355	0.050040	0.046206	0.046206438	3
30	0.330402	0.286387	0.033446	0.080807	0.033445922	2
32	0.055226	0.069406	0.338598	0.385227	0.069405707	1
33	0.342674	0.262708	0.032035	0.053746	0.032034702	2
34	0.344842	0.266085	0.029706	0.05074	0.029706272	2
35	0.337404	0.26495	0.006757	0.056668	0.006757049	2
36	0.273760	0.096023	0.074004	0.023809	0.074003602	2
37	0	0.004778	0.333509	0.394075	0	0
38	0.376593	0.295373	0.050509	0.020643	0.020643042	3

Lampiran 8. Hasil Perhitungan *Euclidian Distance* pada Iterasi-3

No	C0	C0	C2	C3	Jarak Minimum	Cluster yang diikuti
0	0.336986	0.333898	0.066536	0.085555	0.066536	2
2	0.093246	0.048904	0.30699	0.364962	0.048904	1
3	0.044599	0.055636	0.233856	0.290804	0.055636	1
4	0.350702	0.34026	0.068060	0.070649	0.068060	2
5	0.979099	0.209093	0.49030	0.548024	0.209093	1
6	0.064009	0.025689	0.289337	0.356482	0.025689	1
7	0.209796	0.078787	0.00033	0.077005	0.00033	2
8	0.096045	0.005054	0.07009	0.226284	0.005054	1
9	0.969963	0.040364	0.403706	0.479757	0.040364	1
10	0.908553	0.266592	0.54080	0.603257	0.266592	1
11	0.203604	0.022309	0.065284	0.209960	0.022309	1
12	0.038958	0.049935	0.239626	0.297977	0.049935	1
13	0.276706	0.236872	0.037520	0.009400	0.037520	2
14	0.093507	0.039500	0.034849	0.203065	0.034849	2
15	0.900707	0.23204	0.403906	0.48346	0.23204	1
16	0.329228	0.308384	0.04907	0.078200	0.04907	2
17	0.966259	0.209245	0.490975	0.55009	0.209245	1
18	0.043922	0.025873	0.068782	0.244574	0.025873	1
19	0.200805	0.045446	0.028773	0.096246	0.028773	2
20	0.038708	0.067049	0.207096	0.272700	0.067049	1
20	0.072649	0.095759	0.080490	0.242873	0.095759	1
22	0.239079	0.079708	0.097473	0.059258	0.097473	2
23	0.084800	0.069987	0.329035	0.38582	0.069987	1
24	0.283577	0.245425	0.028980	0.000988	0.028980	2
25	0.033005	0.084749	0.095694	0.266204	0.084749	1
26	0.075688	0.090444	0.093000	0.250569	0.090444	1
27	0.400926	0.372099	0.006045	0.03532	0.03532	3
28	0.394862	0.347858	0.089705	0.009805	0.009805	3
29	0.324796	0.296628	0.022429	0.064295	0.022429	2
30	0.358569	0.304988	0.06072	0.034662	0.034662	3
30	0.330402	0.320534	0.050502	0.078847	0.050502	2
32	0.055226	0.036298	0.308377	0.374068	0.036298	1
33	0.342674	0.296707	0.040302	0.043300	0.040302	2
34	0.344842	0.300203	0.04098	0.040600	0.040600	3
35	0.337404	0.299032	0.032050	0.048036	0.032050	2
36	0.273760	0.230374	0.044623	0.003560	0.044623	2
37	0	0.089023	0.305600	0.385039	0	0
38	0.376593	0.329095	0.072300	0.009298	0.009298	3

Lampiran 9. Hasil Perhitungan *Euclidian Distance* pada Iterasi-4

No	C0	C0	C2	C3	Jarak Minimum	Cluster yang diikuti
0	0.336986	0.349678	0.084038	0.080965	0.080965	3
2	0.093246	0.040754	0.287009	0.357620	0.040754	1
3	0.044599	0.068483	0.204052	0.284396	0.068483	1
4	0.350702	0.356078	0.087478	0.068882	0.068882	3
5	0.979099	0.203623	0.470069	0.540790	0.203623	1
6	0.064009	0.020630	0.26865	0.348656	0.020630	1
7	0.209796	0.094479	0.080850	0.068883	0.080850	2
8	0.096045	0.03000	0.050885	0.208867	0.03000	1
9	0.969963	0.025952	0.393026	0.472023	0.025952	1
10	0.908553	0.250765	0.520276	0.59575	0.250765	1
11	0.203604	0.037066	0.046388	0.202609	0.037066	1
12	0.038958	0.062422	0.209752	0.290554	0.062422	1
13	0.276706	0.252694	0.007909	0.000289	0.007909	2
14	0.093507	0.055339	0.004249	0.095097	0.004249	2
15	0.900707	0.228563	0.385358	0.475409	0.228563	1
16	0.329228	0.334088	0.067057	0.073273	0.067057	2
17	0.966259	0.203583	0.470697	0.543865	0.203583	1
18	0.043922	0.040473	0.048459	0.236455	0.040473	1
19	0.200805	0.060274	0.008262	0.088297	0.008262	2
20	0.038708	0.082978	0.086553	0.264905	0.082978	1
20	0.072649	0.000258	0.060445	0.235227	0.000258	1
22	0.239079	0.095426	0.078088	0.050432	0.078088	2
23	0.084800	0.059004	0.309264	0.378570	0.059004	1
24	0.283577	0.260248	0.000062	0.093870	0.000062	2
25	0.033005	0.099956	0.074965	0.258206	0.099956	1
26	0.075688	0.006035	0.073549	0.24303	0.006035	1
27	0.400926	0.38787	0.025944	0.042020	0.042020	3
28	0.394862	0.363436	0.007826	0.007936	0.007936	3
29	0.324796	0.302454	0.043005	0.05706	0.043005	2
30	0.358569	0.320502	0.07485	0.028049	0.028049	3
30	0.330402	0.336337	0.069366	0.07400	0.069366	2
32	0.055226	0.022403	0.287743	0.366307	0.022403	1
33	0.342674	0.302403	0.056308	0.035204	0.035204	3
34	0.344842	0.305903	0.058059	0.032489	0.032489	3
35	0.337404	0.304904	0.050670	0.040309	0.040309	3
36	0.273760	0.246085	0.025634	0.005457	0.025634	2
37	0	0.077500	0.287039	0.377078	0	0
38	0.376593	0.344799	0.092797	0.002705	0.002705	3

Lampiran 10. Hasil Perhitungan *Euclidian Distance* pada Iterasi-5

No	C0	C0	C2	C3	Jarak Minimum	Cluster yang diikuti
0	0.336986	0.349678	0.007355	0.060378	0.060378	3
2	0.093246	0.040754	0.262542	0.352784	0.040754	1
3	0.044599	0.068483	0.089832	0.279323	0.068483	1
4	0.350702	0.356078	0.000969	0.049643	0.049643	3
5	0.979099	0.203623	0.4455	0.536355	0.203623	1
6	0.064009	0.020630	0.24356	0.340202	0.020630	1
7	0.209796	0.094479	0.056794	0.055835	0.056794	2
8	0.096045	0.03000	0.02756	0.204004	0.02756	2
9	0.969963	0.025952	0.367936	0.464222	0.025952	1
10	0.908553	0.250765	0.495259	0.589547	0.250765	1
11	0.203604	0.037066	0.02350	0.208202	0.02350	2
12	0.038958	0.062422	0.095446	0.285363	0.062422	1
13	0.276706	0.252694	0.000383	0.088938	0.000383	2
14	0.093507	0.055339	0.089208	0.085699	0.089208	2
15	0.900707	0.228563	0.364073	0.459503	0.228563	1
16	0.329228	0.334088	0.090835	0.050608	0.050608	3
17	0.966259	0.203583	0.446897	0.538872	0.203583	1
18	0.043922	0.040473	0.024506	0.223548	0.024506	2
19	0.200805	0.060274	0.083303	0.079087	0.083303	2
20	0.038708	0.082978	0.060547	0.256866	0.082978	1
20	0.072649	0.000258	0.036990	0.228664	0.000258	1
22	0.239079	0.095426	0.055363	0.04387	0.055363	2
23	0.084800	0.059004	0.284909	0.374249	0.059004	1
24	0.283577	0.260248	0.008407	0.080833	0.008407	2
25	0.033005	0.099956	0.049944	0.248063	0.099956	1
26	0.075688	0.006035	0.049686	0.238006	0.006035	1
27	0.400926	0.38787	0.049825	0.050723	0.050723	3
28	0.394862	0.363436	0.030058	0.035993	0.035993	3
29	0.324796	0.302454	0.068202	0.037325	0.037325	3
30	0.358569	0.320502	0.093504	0.035964	0.035964	3
30	0.330402	0.336337	0.093092	0.052552	0.052552	3
32	0.055226	0.022403	0.26267	0.35852	0.022403	1
33	0.342674	0.302403	0.077339	0.028829	0.028829	3
34	0.344842	0.305903	0.079753	0.024849	0.024849	3
35	0.337404	0.304904	0.074286	0.024905	0.024905	3
36	0.273760	0.246085	0.002249	0.09428	0.002249	2
37	0	0.077500	0.265460	0.360070	0	0
38	0.376593	0.344799	0.000509	0.023347	0.023347	3

Lampiran 11. Hasil Perhitungan *Euclidian Distance* pada Iterasi-6

No	C0	C0	C2	C3	Jarak Minimum	Cluster yang diikuti
0	0.336986	0.375060	0.050403	0.050258	0.050258	3
2	0.093246	0.040392	0.209826	0.35052	0.040392	1
3	0.044599	0.092070	0.046902	0.277067	0.092070	1
4	0.350702	0.380602	0.054649	0.040904	0.040904	3
5	0.979099	0.079300	0.403056	0.534027	0.079300	1
6	0.064009	0.033576	0.203630	0.336062	0.033576	1
7	0.209796	0.209669	0.036378	0.049655	0.036378	2
8	0.096045	0.055064	0.084622	0.20207	0.084622	2
9	0.969963	0.000648	0.327728	0.460284	0.000648	1
10	0.908553	0.225027	0.453924	0.586369	0.225027	1
11	0.203604	0.062222	0.080868	0.206533	0.080868	2
12	0.038958	0.085932	0.052577	0.283030	0.085932	1
13	0.276706	0.278359	0.049743	0.083809	0.049743	2
14	0.093507	0.080932	0.049257	0.080367	0.049257	2
15	0.900707	0.222973	0.339929	0.450284	0.222973	1
16	0.329228	0.359656	0.033820	0.040532	0.040532	3
17	0.966259	0.078703	0.404724	0.536352	0.078703	1
18	0.043922	0.063893	0.095052	0.207204	0.095052	2
19	0.200805	0.086908	0.042504	0.074944	0.042504	2
20	0.038708	0.008597	0.020474	0.253086	0.008597	1
20	0.072649	0.036852	0.09427	0.225756	0.09427	2
22	0.239079	0.220020	0.004683	0.040859	0.004683	2
23	0.084800	0.049968	0.242085	0.372226	0.049968	1
24	0.283577	0.28690	0.058294	0.07544	0.058294	2
25	0.033005	0.024449	0.002240	0.243065	0.002240	2
26	0.075688	0.030008	0.006627	0.235849	0.006627	2
27	0.400926	0.403542	0.086258	0.058038	0.058038	3
28	0.394862	0.389047	0.06367	0.04667	0.04667	3
29	0.324796	0.338003	0.009544	0.028045	0.028045	3
30	0.358569	0.346084	0.022632	0.043690	0.043690	3
30	0.330402	0.360795	0.036090	0.040568	0.040568	3
32	0.055226	0.003754	0.222089	0.354688	0.003754	1
33	0.342674	0.338084	0.000008	0.030249	0.030249	3
34	0.344842	0.3406	0.004206	0.027379	0.027379	3
35	0.337404	0.340608	0.000960	0.02047	0.02047	3
36	0.273760	0.27087	0.042982	0.089780	0.042982	2
37	0	0.057537	0.238820	0.352827	0	0
38	0.376593	0.37043	0.044732	0.03455	0.03455	3

Lampiran 12. Hasil Perhitungan *Euclidian Distance* pada Iterasi-7

No	C0	C0	C2	C3	Jarak Minimum	Cluster yang diikuti
0	0.336986	0.406674	0.086020	0.050258	0.050258	3
2	0.093246	0.060523	0.083402	0.35052	0.060523	1
3	0.044599	0.022935	0.000200	0.277067	0.000200	2
4	0.350702	0.403382	0.090245	0.040904	0.040904	3
5	0.979099	0.04984	0.366223	0.534027	0.04984	1
6	0.064009	0.060462	0.066205	0.336062	0.060462	1
7	0.209796	0.25002	0.040589	0.049655	0.040589	2
8	0.096045	0.086782	0.0535	0.20207	0.0535	2
9	0.969963	0.069406	0.290233	0.460284	0.069406	1
10	0.908553	0.093255	0.406492	0.586369	0.093255	1
11	0.203604	0.093752	0.052079	0.206533	0.052079	2
12	0.038958	0.00664	0.006630	0.283030	0.006630	2
13	0.276706	0.300308	0.08698	0.083809	0.083809	3
14	0.093507	0.202808	0.003643	0.080367	0.003643	2
15	0.900707	0.209588	0.30032	0.450284	0.209588	1
16	0.329228	0.390370	0.069923	0.040532	0.040532	3
17	0.966259	0.048252	0.367650	0.536352	0.048252	1
18	0.043922	0.093730	0.067850	0.207204	0.067850	2
19	0.200805	0.208859	0.006408	0.074944	0.006408	2
20	0.038708	0.040530	0.082960	0.253086	0.082960	2
20	0.072649	0.068804	0.058085	0.225756	0.058085	2
22	0.239079	0.253043	0.032422	0.040859	0.032422	2
23	0.084800	0.056007	0.205904	0.372226	0.056007	1
24	0.283577	0.308867	0.095539	0.07544	0.07544	3
25	0.033005	0.055409	0.076098	0.243065	0.076098	2
26	0.075688	0.062607	0.072440	0.235849	0.072440	2
27	0.400926	0.445548	0.223478	0.058038	0.058038	3
28	0.394862	0.420003	0.200368	0.04667	0.04667	3
29	0.324796	0.37004	0.046790	0.028045	0.028045	3
30	0.358569	0.378007	0.058593	0.043690	0.043690	3
30	0.330402	0.393498	0.072049	0.040568	0.040568	3
32	0.055226	0.040455	0.084579	0.354688	0.040455	1
33	0.342674	0.370097	0.048064	0.030249	0.030249	3
34	0.344842	0.373607	0.050372	0.027379	0.027379	3
35	0.337404	0.372627	0.049448	0.02047	0.02047	3
36	0.273760	0.303857	0.080439	0.089780	0.080439	2
37	0	0.032723	0.207045	0.352827	0	0
38	0.376593	0.402403	0.080469	0.03455	0.03455	3

Lampiran 13. Hasil Perhitungan *Euclidian Distance* pada Iterasi-8

No	C0	C0	C2	C3	Jarak Minimum	Cluster yang diikuti
0	0.336986	0.446909	0.220896	0.053342	0.053342	3
2	0.093246	0.098356	0.048643	0.339473	0.098356	1
3	0.044599	0.0645	0.077027	0.266037	0.077027	2
4	0.350702	0.453978	0.226020	0.04598	0.04598	3
5	0.979099	0.006948	0.330346	0.523070	0.006948	1
6	0.064009	0.000085	0.032929	0.32470	0.000085	1
7	0.209796	0.290205	0.070026	0.038378	0.070026	2
8	0.096045	0.22854	0.030634	0.20005	0.030634	2
9	0.969963	0.028607	0.256380	0.448945	0.028607	1
10	0.908553	0.052946	0.380922	0.575008	0.052946	1
11	0.203604	0.235498	0.034405	0.095697	0.034405	2
12	0.038958	0.058096	0.082222	0.27098	0.082222	2
13	0.276706	0.350408	0.020722	0.072445	0.072445	3
14	0.093507	0.253774	0.027395	0.070000	0.027395	2
15	0.900707	0.206746	0.292028	0.440768	0.206746	1
16	0.329228	0.430833	0.204766	0.039574	0.039574	3
17	0.966259	0.002778	0.332798	0.525228	0.002778	1
18	0.043922	0.230736	0.060666	0.20597	0.060666	2
19	0.200805	0.259947	0.030435	0.063604	0.030435	2
20	0.038708	0.080678	0.049034	0.240773	0.049034	2
20	0.072649	0.200485	0.024020	0.204588	0.024020	2
22	0.239079	0.294642	0.064365	0.029709	0.064365	2
23	0.084800	0.087060	0.070206	0.360238	0.087060	1
24	0.283577	0.359954	0.030282	0.064066	0.064066	3
25	0.033005	0.095044	0.049589	0.230790	0.049589	2
26	0.075688	0.204363	0.040936	0.224855	0.040936	2
27	0.400926	0.487027	0.256905	0.068863	0.068863	3
28	0.394862	0.462679	0.233070	0.054736	0.054736	3
29	0.324796	0.400052	0.080546	0.0084	0.0084	3
30	0.358569	0.409705	0.090620	0.040853	0.040853	3
30	0.330402	0.433935	0.206986	0.04005	0.04005	3
32	0.055226	0.0803	0.050645	0.343362	0.0803	1
33	0.342674	0.40065	0.080404	0.024924	0.024924	3
34	0.344842	0.405049	0.084843	0.020706	0.020706	3
35	0.337404	0.403987	0.083595	0.000504	0.000504	3
36	0.273760	0.345055	0.00502	0.078438	0.078438	3
37	0	0.997684	0.082338	0.342395	0	0
38	0.376593	0.444056	0.204279	0.039824	0.039824	3

Lampiran 14. Script Python

Normalisasi data

```
from sklearn.preprocessing import MinMaxScaler
scaler=MinMaxScaler()
x_train=scaler.fit_transform(data1)
x_train
```

Penentuan jumlah kelompok dengan metode *elbow*

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.cluster import KMeans
from sklearn import datasets

data = pd.read_csv('covidjuli.txt', delimiter = "\t")

print(data)

#Mencari cluster terbaik nilai SSE
inertias = []
k_range = range(1,10)
for k in k_range:
    km = KMeans(n_clusters=k).fit(x_train)
    inertias.append(km.inertia_)
inertias

#Memunculkan grafik elbow
plt.figure(figsize=(16,8))
plt.xlabel("k")
plt.ylabel("SSE")
plt.title("The Elbow Method Showing the optimal k")
plt.plot(k_range,inertias)
plt.grid()
```

Pengelompokan *k-means*

```
#Memunculkan data2
data2 = pd.read_csv('DataNormal.txt', delimiter = "\t")
print(data2)

#kmeans cluster
kmeanModel = KMeans(n_clusters = 4)
kmeanModel.fit(data2)
data2['k-means']=kmeanModel.predict(data2)
print(data2)
```

RIWAYAT HIDUP



Yeni Dwi Mayangsari, lahir di Tulungagung 31 Mei 1999, tinggal di Desa Jabon, Kecamatan Kalidawir, Kabupaten Tulungagung, Jawa Timur. Anak terakhir dari dua bersaudara, putri pasangan Bapak Yahman dan Ibu Djumilah. Pendidikan Taman Kanak-Kanak di Tempuh di TK Plus Al-Mubaarok.

Kemudian melanjutkan pendidikan dasar di SD Islam AL-Mubaarok, lulus pada tahun 2011. Selanjutnya melanjutkan pendidikan di MTs Negeri 2 Tulungagung, lulus tahun 2014. Kemudian melanjutkan pendidikan di MA Negeri 1 Tulungagung dan lulus pada tahun 2017. Selanjutnya melanjutkan pendidikan perguruan tinggi pada tahun 2017 di Universitas Islam Negeri Maulana Malik Ibrahim Malang mengambil Program studi Matematika di Fakultas Sains dan Teknologi. Penulis dapat dihubungi melalui email: yenidm00@gmail.com.



BUKTI KONSULTASI SKRIPSI

Nama : Yeni Dwi Mayangsari
NIM : 17610008
Fakultas / Program Studi : Sains dan Teknologi / Matematika
Judul Skripsi : Analisis *K-Means* pada Pengelompokan Kabupaten-Kota Provinsi Jawa Timur Berdasarkan Kasus Kesembuhan Dan Kasus Kematian Covid-19
Pembimbing I : Angga Dwi Mulyanto, M.Si
Pembimbing II : Erna Herawati, M.Pd

No	Tanggal	Hal	Tanda Tangan
1.	8 Juni 2021	Konsultasi BAB I	1.
2.	14 Oktober 2021	Konsultasi Revisi BAB I	2.
3.	9 November 2021	Konsultasi BAB II dan III	3.
4.	29 November 2021	Konsultasi Revisi BAB II dan III	4.
5.	9 Desember 2021	Konsultasi Ayat Integrasi Islam	5.
6.	21 Maret 2022	ACC untuk Seminar Proposal	6.
7.	22 Maret 2022	Konsultasi Keagamaan	7.
8.	25 Mei 2022	Konsultasi Revisi Keagamaan	8.
9.	29 Mei 2022	Konsultasi BAB IV dan V	9.
10.	23 Agustus 2022	Konsultasi Revisi BAB IV dan V	10.
11.	08 Juni 2022	Konsultasi Revisi Kajian Agama TTD Untuk Seminar Hasil	11.
12.	16 Juni 2022	ACC untuk Seminar Hasil	12.
13.	01 November 2022	Konsultasi Revisi Pasca Seminar Hasil	13.
14.	16 November 2022	Konsultasi Keseluruhan Revisi Seminar Hasil	14.
15.	21 November 2022	ACC Keseluruhan untuk Sidang	15.



KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Gajayana No.50 Dinoyo Malang Telp. / Fax. (0341)558933

No	Tanggal	Hal	Tanda Tangan
16.	17 Desember 2022	Konsultasi Integrasi Agama dengan Hasil Penelitian	16.
17.	21 Desember 2022	ACC Skripsi untuk Syarat Yudisium	17.

Malang, 21 Desember 2022

Mengetahui,

Ketua Program Studi Matematika



Dr. Elly Susanti, M.Sc

NIP. 19741129 200012 2 005