

**PENGELOMPOKAN DAERAH TERDAMPAK COVID-19 DI PROVINSI
JAWA TIMUR MENGGUNAKAN METODE *HIERARCHICAL K-MEANS*
*CLUSTERING (HK-MEANS)***

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh
MU'ASATUDZ DZIKRIYYAH
H72217053

**PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2021

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : MU'ASATUDZ DZIKRIYYAH

NIM : H72217053

Program Studi : Matematika

Angkatan : 2017

Mengatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "PENGELOMPOKAN DAERAH TERDAMPAK COVID-19 DI PROVINSI JAWA TIMUR MENGGUNAKAN METODE *HIERARCHICAL K-MEANS CLUSTERING (HK-MEANS)*". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian inisaya buat dengan sebenar-benarnya.

Surabaya, 17 Juni 2021

Yang menyatakan,



MU'ASATUDZ DZIKRIYYAH

NIM.H72217053

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : MU'ASATUDZ DZIKRIYYAH
NIM : H72217053
Judul Skripsi : PENGELOMPOKAN DAERAH TERDAMPAK COVID-19 DI PROVINSI JAWA TIMUR MENGGUNAKAN METODE *HIERARCHICAL K-MEANS CLUSTERING (HK-MEANS)*

telah diperiksa dan disetujui untuk diujikan.

Surabaya, 17 Juni 2021

Pembimbing I



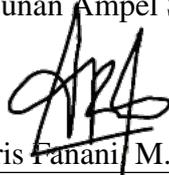
Nurissaidah Minnuha, M.Kom
NIP. 199011022014032004

Pembimbing II



Lutfi Hakim, M.Ag
NIP. 197312252006041001

Mengetahui,
Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya



Aris Fanani, M.Kom
NIP. 198701272014031002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

Nama : MU'ASATUDZ DZIKRIYYAH
NIM : H72217053
Judul Skripsi : PENGELOMPOKAN DAERAH TERDAMPAK COVID-19 DI PROVINSI JAWA TIMUR MENGGUNAKAN METODE *HIERARCHICAL K-MEANS CLUSTERING (HK-MEANS)*

Telah dipertahankan di depan Tim Penguji
pada tanggal 17 Juni 2021

Mengesahkan,
Tim Penguji

Penguji I

Nurissaidah Ulimnuha, M.Kom
NIP. 199011022014032004

Penguji II

Lutfi Hakim, M.Ag
NIP. 197312252006041001

Penguji III

Dian Candra Rini Novitasari, M.Kom
NIP. 198511242014032001

Penguji IV

Yuniar Farida, ST
NIP. 19790527014032002

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya



Dr. Hj. Ev. Ratimatur Rusydiyah, M.Ag
NIP. 197312272005012003



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : MU'ASATUD2 DZIKRIYYAH
NIM : H72217053
Fakultas/Jurusan : SAINTEK / MATEMATIKA
E-mail address : muasatud2d2gb@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Sekripsi Tesis Desertasi Lain-lain (.....)
yang berjudul :

PENGELOMPOKAN DAERAH TERDAMPAK COVID-19 DI PROVINSI
JAWA TIMUR MENGGUNAKAN METODE HIERARCHICAL K-MEANS
CLUSTERING (HK-MEANS)

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 16 Juli 2021

Penulis

(MU'ASATUD2 D.)

2.6.1. Metode <i>agglomerative</i> (penggabungan)	18
2.6.2. Metode <i>Divisive</i> (pemecahan)	20
2.7. <i>K-Means</i>	20
2.8. Metode <i>Hierarchical K-Means Clustering</i>	22
2.9. Pengujian Metode Terbaik dengan <i>Silhouette coefficient</i>	24
2.10. Kewajiban Menjaga Kesehatan dan Keselamatan Jiwa	26
2.10.1. Memakai Masker	27
2.10.2. Selalu Mencuci tangan	28
2.10.3. Tetap berada di rumah	29
2.10.4. Tidak Berprasangka Buruk Kepada Allah Ta'ala	30
2.10.5. Bersikap Optimis dan Berucap yang Baik	31
III METODE PENELITIAN	32
3.1. Jenis Penelitian	32
3.2. Sumber Data	32
3.3. Variabel Penelitian	32
3.4. Tahapan Penelitian	33
IV HASIL DAN PEMBAHASAN	37
4.1. Deskripsi Data	37
4.2. Standarisasi Data	39
4.3. Uji Multikolinieritas	39
4.4. Proses <i>Hierarchical K-Means Clustering</i>	44
4.4.1. <i>Hierarchical Clustering</i>	44
4.4.2. Menentukan Pusat <i>Cluster (Centroid)</i>	54
4.4.3. <i>K-Means</i>	57
4.5. Pemilihan Metode Terbaik Dengan <i>Sillhoutte Coefficient</i>	61
4.6. Interpretasi <i>Cluster Optimal</i>	64
4.7. Pemetaan Daerah Terdampak Covid 19 di Jawa Timur Bulan April-Desember 2020	66
4.8. Integrasi Keislaman	74
V PENUTUP	77

5.1. Simpulan	77
5.2. Saran	78
DAFTAR PUSTAKA	79
A Data Kasus Covid-19 di Jawa Timur Bulan April 2020	84
B Data Kasus Covid-19 di Jawa Timur Bulan Mei 2020	85
C Data Kasus Covid-19 di Jawa Timur Bulan Juni 2020	86
D Data Kasus Covid-19 di Jawa Timur Bulan Juli 2020	87
E Data Kasus Covid-19 di Jawa Timur Bulan Agustus 2020	88
F Data Kasus Covid-19 di Jawa Timur Bulan September 2020	89
G Data Kasus Covid-19 di Jawa Timur Bulan Oktober 2020	90
H Data Kasus Covid-19 di Jawa Timur Bulan Nopember 2020	91
I Data Kasus Covid-19 di Jawa Timur Bulan Desember 2020	92
J TABEL STANDARISASI DATA Z-SCORE	93
K TABEL JARAK EUCLIDEAN	94
L Jumlah <i>Cluster</i> dan Anggota Metode <i>Single Linkage</i> Bulan Desember 2020	100
M Jumlah <i>Cluster</i> dan Anggota Metode <i>Complete Linkage</i> Bulan Desember 2020	101
N Jumlah <i>Cluster</i> dan Anggota Metode <i>Average Linkage</i> Bulan Desember 2020	102
O Jumlah <i>Cluster</i> dan Anggota Metode <i>HK-Means Single Linkage</i> Bulan Desember 2020	103
P Jumlah <i>Cluster</i> dan Anggota Metode <i>HK-Means Complete Linkage</i> Bulan Desember 2020	104
Q Jumlah <i>Cluster</i> dan Anggota Metode <i>HK-Means Average Linkage</i> Bulan Desember 2020	105

bahwa: barangsiapa yang membunuh seorang manusia, bukan karena orang itu (membunuh) orang lain, atau bukan karena membuat kerusakan dimuka bumi, maka seakan-akan dia telah membunuh manusia seluruhnya. Dan barangsiapa yang memelihara kehidupan seorang manusia, maka seolah-olah dia telah memelihara kehidupan manusia semuanya. Dan sesungguhnya telah datang kepada mereka rasul-rasul Kami dengan (membawa) keterangan-keterangan yang jelas, kemudian banyak di antara mereka sesudah itu sungguh-sungguh melampaui batas dalam berbuat kerusakan dimuka bumi. (QS Al-Maidah :32)

Pada 12 Maret 2020 WHO (*World Health Organization*) mengumumkan bahwa Covid-19 sebagai pandemik. Virus ini sudah menyebar di 215 negara di dunia termasuk Indonesia. Berdasarkan informasi yang dipublikasikan oleh Kementerian Kesehatan Republik Indonesia sejak pertama kali dilaporkan pada bulan Desember 2019 sampai tanggal 31 Desember 2020, di Indonesia telah terkonfirmasi kasus Covid-19 sebanyak 743.198 kasus dengan jumlah kematian sebanyak 22.138 kasus.

Khususnya di provinsi Jawa Timur yang merupakan salah satu wilayah di Indonesia yang penyebaran Covid-19 juga terus meningkat dan belum menunjukkan penurunan sampai saat ini. Pada tanggal yang sama juga diinformasikan pada laman web yang dipublikasikan oleh Kementerian Kesehatan Republik Indonesia bahwa jumlah kasus positif Covid-19 di Jawa Timur mengalami penambahan kasus baru sebanyak 935 kasus baru jadi total sebanyak 84.152 kasus dan jumlah kasus yang sembuh sebanyak 72.135 kasus dengan jumlah kematian sebanyak 5.827 kasus. Dengan penambahan kasus tersebut membuat Jawa Timur menduduki peringkat empat kasus Covid-19 terbanyak setelah DKI Jakarta dan Jawa Tengah. Bahkan Jawa timur sempat menempati peringkat pertama kasus covid-19 terbanyak yaitu

pada 28 Juni 2020 dengan penambahan kasus baru sebanyak 330 kasus.

Tingginya penyebaran Covid-19 di Jawa Timur ini maka perlu adanya pengelompokan daerah terdampak Covid-19 di Jawa Timur untuk mengetahui kemiripan karakteristik atau kriteria dari setiap daerah. Daerah yang memiliki karakteristik yang sama atau hampir sama akan terkumpul dalam suatu kelompok. Hal ini bertujuan untuk memberikan informasi terkait Covid-19 kepada masing-masing daerah. Salah satu cara yang dapat dilakukan dalam pengelompokan ini yaitu dengan analisis *cluster*.

Pada penelitian sebelumnya yaitu penelitian yang dilakukan oleh Rini Silvi (2018) yang membandingkan metode *centroid linkage* dan *K-means* dalam menentukan hasil analisis *cluster* dengan data outlier untuk pengelompokan indikator HIV/AIDS di Indonesia. Penelitian ini menggunakan gap statistik untuk menentukan jumlah *cluster* ideal yang mengelompokkan propinsi berdasarkan indikator HIV/AIDS sedemikian hingga terbagi menjadi 7 *cluster*. Hasil penelitian ini memberikan kesimpulan bahwa untuk data yang memiliki *outlier*, metode pengclusteran menggunakan *centroid linkage* lebih memberikan hasil yang sesuai dibandingkan dengan metode *K-means*. Rasio S_w/S_b metode *K-means* adalah sebesar 0,112232 dan rasio S_w/S_b metode *centroid linkage* menghasilkan angka 0,067307. Metode *centroid linkage* menghasilkan kelompok yang lebih homogen sehingga nilai rasio yang dihasilkan lebih kecil. Artinya metode *centroid linkage* memiliki kualitas ketepatan kelompok yang lebih baik dibandingkan metode *K-means*. (Silvi, 2018)

Selain itu penelitian lainnya juga dilakukan oleh Lynda Rahmawati, Sari Widyia Sihwi dan Esti Suryani (2016) dengan judul *Analisa clustering* menggunakan metode *K-means* dan *hierarchical clustering* kasus : dokumen skripsi jurusan kimia, FMIPA Universitas Sebelas Maret. Pengelompokan data penelitian yang ber-

bentuk teks dapat dilakukan dengan *text mining* dengan metode *clustering*. Metode *clustering* yang digunakan adalah kombinasi antara metode *K-means* dan *hierarchical clustering*. Data penelitian yang dipilih adalah dokumen skripsi bagian abstrak. *Clustering* dokumen menghasilkan 16 *cluster*. Hasil analisa *cluster* memperlihatkan bahwa keahlian dosen mempengaruhi variasi tema penelitian yang dilakukan oleh mahasiswa. (Rahmawati et.al , 2016)

Analisa Perbandingan Metode *Hierarchical Clustering*, *K-means* dan Gabungan Keduanya dalam *Cluster Data* (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS) yang dilakukan oleh Tahta Alfina, Budi Santosa, dan Ali Ridho Barakbah. *Clustering* memiliki dua metode, yaitu partisi dan hierarki. Dua metode ini memiliki kelebihan dan kekurangan masing-masing, dan dengan menggabungkan keduanya dapat diperoleh hasil *cluster* yang lebih baik. Dari hasil *cluster* dengan menggunakan data problem Kerja Praktek Jurusan Teknik Industri ITS, maka diperoleh hasil bahwa kombinasi algoritma *hierarchical clustering* dan *k-means* menghasilkan pengelompokan data yang lebih baik. Pada skenario jumlah *cluster* sebanyak 5, penjumlahan nilai s terbesar diperoleh oleh metode *single linkage clustering* yang dikombinasikan dengan *k-means* sebesar 300 data diclusterkan pada *cluster* yang tepat dengan nilai koefisien kolerasi sebesar 0,9118, diikuti oleh 3 metode *hierarchical clustering* yang lainnya yang digabungkan dengan *k-means* dan penjumlahan nilai s yang paling kecil dihasilkan oleh metode *k-means* yaitu sebesar 210 data. (Alfina et.al , 2019)

Penelitian yang lain dilakukan oleh Ali Mahmudan (2020) dengan Pengelompokan Kabupaten/Kota di Jawa Tengah Berdasarkan Kasus COVID-19 Menggunakan *K-Means Clustering*. Berdasarkan penelitian tersebut menunjukkan hasil Pengelompokan kabupaten/kota di Provinsi Jawa Tengah berdasarkan kasus COVID-

19 menghasilkan sejumlah 33 kabupaten/kota berada pada cluster 1, 1 kota berada pada cluster 2, dan 1 kabupaten berada pada clusters 3. Kondisi kasus COVID-19 di Provinsi Jawa Tengah sudah relatif aman tetapi harus tetap dilakukan penanganan semaksimal mungkin agar kasus aktif dan kasus meninggal tidak semakin banyak. Selain itu, cluster 2 yaitu Kota Semarang dan cluster 3 yaitu Kabupaten Jepara harus lebih diperhatikan oleh Pemerintah Provinsi Jawa Tengah karena wilayah tersebut masih memiliki jumlah kasus aktif yang tinggi. Hal tersebut harus dilakukan agar tidak terjadi penambahan kasus meninggal maupun kasus aktif di wilayah tersebut atau penyebaran kasus aktif di luar wilayah tersebut. (Mahmudah, 2020)

Berdasarkan pada penjelasan tersebut, penulis tertarik dalam melakukan penelitian mengenai analisis cluster dengan mengambil kasus Covid-19 yang sekarang menjadi pandemi di seluruh dunia. Judul dalam penelitian ini adalah "Pengelompokan Daerah Terdampak Covid-19 Di Provinsi Jawa Timur Menggunakan Metode *Hierarchical K-Means Clustering (HK-Means)* (Studi Kasus : Jawa Timur)". Metode yang digunakan tersebut merupakan gabungan antara metode *hierarchical clustering* dan *K-means*, digabungkan agar dapat meningkatkan dan menghasilkan hasil *cluster* dengan kualitas yang lebih baik dan tinggi dan pada berdasarkan penelitian alfinia terbukti bahwa gabungan metode tersebut menghasilkan hasil yang lebih baik dibandingkan dengan metode *K-means*. Karena hasil akhir pengelompokan dengan metode *K-means* sangat sensitif terhadap pemilihan pusat *cluster* jadi hasilnya sedikit berbeda setiap kali dihitung. Metode *K-means* mempunyai kelemahan yaitu pada penentuan pusat awal *cluster*. Hasil *cluster* yang terbentuk dari metode *K-Means* ini sangatlah tergantung pada inisiasi nilai pusat awal *cluster* yang diberikan. Untuk itu, metode *K-means* digabungkan dengan metode *Hierarchical Clustering* untuk penentuan pusat awal *cluster* untuk menghasilkan nilai yang le-

yang terdapat dalam proses *clustering* dinamakan dengan *cluster*. *Cluster* merupakan suatu kelompok data yang mempunyai tingkat kesamaan dengan suatu *cluster* dan karakteristiknya tidak sama dengan *cluster* lainnya. (Febritanti et.all , 2016)

Larose dalam buku yang ditulis oleh Kusrini dan Luthfi mengelompokan Data Mining dapat dibagi menjadi 6 kelompok yaitu deskripsi, estimasi, prediksi, klasifikasi, *clustering* (pengelompokan), dan asosiasi. Berbeda dengan klasifikasi, teknik *clustering* yaitu mengelompokkan data secara otomatis sebelum mengetahui label kelasnya. Tetapi *clustering* juga dapat digunakan untuk himpunan data dengan label kelas yang diketahui (Suryanto , 2017).

Metode *clustering* sering digunakan sebagai dasar penentuan kelas untuk data-data yang kelasnya belum ditentukan. Dasar dari pengonsepan *clustering* adalah melakukan pengelompokkan beberapa data ke dalam suatu *cluster* yang sesuai yaitu *cluster* yang mempunyai tingkat ketidaksamaan yang tinggi terhadap objek *cluster* lainnya. (Abdillah et.all , 2016)

Sehingga, dari semua pengertian bahwa *clustering* adalah suatu metode yang digunakan untuk memberikan label bagi data yang belum diketahui secara pasti kelas *cluster*. Ciri-ciri *cluster* adalah sebagai berikut (Santoso dan Slamet , 1987) :

1. Homogenitas (kesamaan) antar anggota *cluster* tinggi (*within cluster*).
2. Heterogenitas (perbedaan) antar *cluster* yang tinggi (*between cluster*).

Menurut (N. Alshatri et.all , 2014), *clustering* dikelompokkan menjadi 5 kelompok besar yaitu :

1. *Partitioning Based Clustering*

Algoritma *Partitional Clustering* membagi titik data ke dalam partisi k , di mana setiap partisi mewakili sebuah *cluster*. Contoh algoritma *partitio-*

ning based clustering adalah *k-means*, *k-medoids*, *k-modes*, *PAM*, *CLARANS*, *FCM* (Chitra dan Maheswari, 2017).

2. Hierarchical Based Clustering

Hierarchical Based clustering adalah metode analisis *cluster* yang digunakan untuk membangun hierarki *cluster*. Dalam data mining, cara kerja *hierarchical clustering* ialah dengan mengelompokkan objek-objek data ke dalam pohon *cluster*, yang mana struktur pohon *cluster* ini biasa disebut dendrogram. Contoh algoritma *hierarchical based clustering* adalah *BIRCH*, *CURE*, *ROCK*, dll (Rani dan Rohil at.all , 2013).

3. Density Based Clustering

Algoritma ini dapat menemukan *cluster* yang berbentuk acak dan dapat menangani *noise*, dan juga merupakan algoritma satu-scan yang mana hanya melakukan satu kali *scanning* ketika memeriksa data mentah. Contoh algoritma *density based clustering* ialah *DBSCAN*, *OPTICS*, *LDBSCAN*, dll (Amini et.all , 2011).

4. Grid Based Clustering

Algoritma *Grid-Based Clustering* adalah tipe yang paling penting dalam algoritma *clustering* hierarki. Pendekatan *Grid-Based* lebih menganggap sel dibandingkan titik data. Karena sifat algoritma *Grid-Based Clustering* ini umumnya lebih efisien secara komputasi apabila dibandingkan dengan semua jenis algoritma *clustering*. Contoh algoritma *grid based clustering* adalah *wave-cluster*, *STING*, *CLIQUE*, *OptiGrid* (N. A. Shah dan Paul , 2017).

5. Model Based Clustering

3. Melakukan uji multikolinearitas terhadap data untuk mengetahui suatu data terjadi korelasi antar variabel atau tidak dengan menggunakan persamaan 2.2
4. Menentukan nilai k . Pada penelitian ini dilakukan ujicoba k dengan $k = 2, 3, 4$, dan 5.
5. Melakukan pengelompokkan dengan metode *Hierarchical clustering* menggunakan teknik pengelompokkan *Agglomeratif* yaitu *Single Linkage*, *Complete Linkage* dan *Average Linkage*.
6. Melakukan pengelompokan setiap daerah dengan metode *Single Linkage*. Tahapan pengelompokkan dengan metode *Single Linkage*, *Complete Linkage*, dan *Complete Linkage* sebagai berikut :
 - (a) Menghitung jarak antar objeknya dengan jarak *euclidean* dengan Persamaan 2.5.
 - (b) Kemudian, dari hasil perhitungan tadi dipilih jarak yang paling minimal atau yang terdekat antar objek.
 - (c) Merhitungan ulang jarak dengan metode *Single Linkage* dengan Persamaan 2.6, metode *Complete Linkage* dengan Persamaan 2.7 dan metode *Average Linkage* dengan Persamaan 2.8.
 - (d) Hal ini akan terus dilakukan dan akan berhenti jika memenuhi kondisi jumlah $k = 1$ atau tersisasa satu *cluster*.
 - (e) Menentukan jumlah anggota *cluster* beserta anggotanya dengan membentuk dendrogram yang sesuai dengan *cluster*.
7. Pada akhir tahap *Hierarchical Clustering* ini akan diperoleh sebuah gambar dendrogram yang menunjukkan urutan pengelompokan masing-masing anggota dalam *cluster*.

Berdasarkan pada Tabel 3.1 terdapat variabel data yang mengalami multikolinieritas yaitu variabel x_1 dan x_2 karena nilai $VIF > 10$ dan nilai $tolerance < 0,10$. Untuk menangani variabel data yang multikolinieritas tersebut yaitu dengan mengeluarkan salah satu variabel agar tidak multikolinieritas. Pada variabel data tersebut yang dikeluarkan adalah variabel x_2 .

4.4. Proses *Hierarchical K-Means Clustering*

Metode *Hierarchical K-Means Clustering* merupakan metode pengelompokan lanjutan, yaitu dengan menggabungkan metode *Hierarchical* dan metode *K-Means non Hierarchical* untuk menghasilkan hasil atau solusi *cluster* yang lebih baik.

4.4.1. *Hierarchical Clustering*

Metode *Hierarchical* yang akan digunakan pada penelitian ini adalah metode *Single Linkage*, *Complete Linkage* dan *Average Linkage*. Ketiga metode ini akan dibandingkan hasilnya untuk mengetahui metode mana yang menghasilkan hasil *cluster* terbaik.

1. *Single Linkage*

Pada metode *Single Linkage* proses pengelompokannya yaitu dengan menggabungkan dua objek atau lebih yang memiliki jarak paling dekat. Untuk mengukur kedekatan atau kemiripan antar dua objek digunakan jarak *Euclidean* menggunakan Persamaan 2.5.

Data yang akan digunakan adalah data yang telah distandarisasi menggunakan *z-score* dan tidak multikolinieritas pada Lampiran J. Dibawah ini merupakan contoh dari perhitungan jarak antara objek 1 (Kab. Situbondo)

$$d_{1,2} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \dots + (x_7 - y_7)^2}$$

$$d_{1,2} = \sqrt{((-0,2831) - 1,60404)^2 + ((-0,1474) - 0,0448)^2}$$

$$\sqrt{+((-0,2031) - 1,656)^2 + \dots + ((0,2261 - 1,4080)^2}$$

$$d_{1,2} = 4,1142566$$

Dari hasil perhitungan jarak *Euclidean* tersebut antara objek 2 (Kab. Madiun) dan objek 3 (Kab. Sumenep) didapatkan jarak sebesar 2,4165608, antara objek 2 (Kab. Madiun) dan objek 3 (Kab. Sumenep) didapatkan jarak sebesar 1,8794918 dan antara objek 3 (Kab. Sumenep) dan objek 4 (Kota Surabaya) didapatkan jarak sebesar 4,1142566. Hal itu menunjukkan bahwa jarak *Euclidean* antara Kab. Madiun dan Kab. Sumenep lebih dekat dan memiliki karakteristik yang lebih mirip dibandingkan dengan Kab. Madiun dan Kab. Situbondo. Hasil perhitungan jarak *Euclidean* antar kabupaten dan kota dapat dilihat pada Lampiran K.

Membentuk *cluster* baru dilakukan dengan cara menggabungkan jarak antara dua objek yang paling dekat atau yang mempunyai jarak terkecil. Berdasarkan Lampiran K, jarak antar kabupaten atau kota yang paling terkecil sebesar 0.39518 yaitu jarak antara *cluster* objek 2 (Kab. Madiun) dan *cluster* objek 32 (Kota Madiun), kemudian kedua objek tersebut menjadi satu *cluster*.

Kemudian dilakukan pembentukan matriks jarak baru dengan menggunakan Persamaan 2.6 berdasarkan Lampiran K. Jarak antar kabupaten atau kabupaten yang digunakan untuk pembentukan *cluster* baru adalah menggunakan objek 2 (Kab. Madiun) dan objek 32 (Kota Madiun). Perhitungan

Kab. Pamekasan	1,565	6,148	1
Kab. Bojonegoro	1,618	6,174	1
Kota Mojokerto	1,828	5,898	1
Kab. Lamongan	1,696	6,659	1
Kab. Trenggalek	1,078	5,709	1
Kab. Pasuruan	1,757	6,162	1
Kota Blitar	1,459	5,911	1
Kab. Pacitan	1,231	5,856	1
Kab. Banyuwangi	5,585	8,585	1
Kab. Kediri	3,614	8,071	1
Kota Kediri	5,920	0,000	2
Kab. Ponorogo	1,329	6,257	1
Kab. Blitar	1,459	6,334	1
Kab. Sampang	2,185	5,487	1
Kota Malang	4,642	8,485	1
Kab. Jember	7,043	10,684	1
Kab. Sidoarjo	1,669	6,402	1
Kab. Tulungagung	1,539	5,009	1
Kota Batu	1,915	5,917	1
Kab. Jombang	2,041	6,609	1
Kab. Ngawi	1,584	5,383	1
Kab. Nganjuk	1,326	5,509	1
Kota Probolinggo	1,589	5,431	1
Kab. Tuban	2,706	7,228	1
Kab. Mojokerto	1,734	5,960	1

Tabel 4.12 Pemetaan Kabupaten/Kota Berdasarkan Hasil Analisis *Cluster*

Kab/Kota	Apr	Mei	Juni	Juli	Agt	Sep	Okt	Nop	Des
Kab. Situbondo	H	H	H	H	H	H	H	H	H
Kab. Madiun	H	H	H	H	H	H	H	H	H
Kab. Sumenep	H	H	H	H	H	H	H	H	H
Kota Surabaya	M	M	M	M	M	M	M	M	M
Kab. Gresik	H	H	H	H	H	H	H	H	H
Kab. Pamekasan	H	H	H	H	H	H	H	H	H
Kab. Bojonegoro	H	H	H	H	H	H	H	H	H
Kota Mojokerto	H	H	H	H	H	H	H	H	H
Kab. Lamongan	H	H	H	H	H	H	H	H	H
Kab. Trenggalek	H	H	H	H	H	H	H	H	H
Kab. Pasuruan	H	H	H	H	H	H	H	H	H
Kota Blitar	H	H	H	H	H	H	H	H	H
Kab. Pacitan	H	H	H	H	H	H	H	H	H
Kab. Banyuwangi	H	H	H	H	H	H	H	M	M

Kab/Kota	Apr	Mei	Juni	Juli	Agt	Sep	Okt	Nop	Des
Kab. Kediri	H	H	H	H	H	H	H	H	M
Kota Kediri	H	H	H	H	H	H	H	H	H
Kab. Ponorogo	H	H	H	H	H	H	H	H	H
Kab. Blitar	H	H	H	H	H	H	H	H	H
Kab. Sampang	H	H	H	H	H	H	H	H	H
Kota Malang	H	H	H	H	H	H	H	H	M
Kab. Jember	H	H	H	H	H	H	H	M	M
Kab. Sidoarjo	H	H	H	H	H	M	M	H	H
Kab. Tulungagung	H	H	H	H	H	H	H	H	H
Kota Batu	H	H	H	H	H	H	H	H	H
Kab. Jombang	H	H	H	H	H	H	H	M	H
Kab. Ngawi	H	H	H	H	H	H	H	H	H
Kab. Nganjuk	H	H	H	H	H	H	H	H	H
Kota Probolinggo	H	H	H	H	H	H	H	H	H
Kab. Tuban	H	H	H	H	H	H	H	H	M
Kab. Mojokerto	H	H	H	H	H	H	H	H	H
Kab. Magetan	H	H	H	H	H	H	H	H	H
Kota Madiun	H	H	H	H	H	H	H	H	H
Kab. Lumajang	H	H	H	H	H	H	H	M	H
Kab. Bondowoso	H	H	H	H	H	H	H	H	H
Kota Pasuruan	H	H	H	H	H	H	H	H	H
Kab. Malang	H	H	H	H	H	H	H	H	H
Kab. Probolinggo	H	H	H	H	H	H	H	H	H
Kab. Bangkalan	H	H	H	H	H	H	H	H	H

Banyuwangi, Kabupaten Jombang, Kabupaten Lumajang dan Kabupaten Jember, Kabupaten Sidoarjo sudah tidak termasuk anggota *cluster* daerah yang rawan lagi dan sudah menjadi anggota *cluster* daerah yang aman. Dan pada bulan Desember 2020 ada penambahan anggota *cluster* lagi daerah yang rawan (Zona merah) Covid-19 yaitu Kabupaten Kediri, Kota Malang, Kabupaten Tuban dan Kabupaten Jember, sedangkan Kabupaten Banyuwangi, Kabupaten Jombang, Kabupaten Lumajang sudah tidak termasuk anggota dari *cluster* daerah rawan dan menjadi anggota *cluster* daerah aman.

Kota Surabaya menjadi satu-satunya yang berada pada daerah yang rawan (Zona merah) Covid-19 dari bulan April-Desember 2020. Hal tersebut dipengaruhi oleh beberapa faktor yaitu, karena kota Surabaya merupakan ibukota provinsi Jawa Timur sehingga penduduknya banyak dan aktivitas masyarakatnya juga tinggi. Beberapa faktor lainnya yaitu karena banyak masyarakatnya yang melakukan aktivitas tanpa mematuhi protokol kesehatan. Dan faktor penyebab lonjakan kasus positif lainnya di Kota Surabaya adalah karena budaya dan karakteristik masyarakatnya, masyarakat Surabaya terkenal dengan sifat yang keras dan berani. Dan pada masyarakat Surabaya terdapat budaya "cangkruan". Cangkruan atau cangkruk adalah kegiatan yang digunakan sebagai tempat pusat komunikasi, sosialisasi, informasi dan hiburan. Budaya cangkruk ini telah menjadi gaya hidup banyak masyarakat Surabaya. Dengan budaya sehingga memunculkan banyak kerumunan yang berdampak pada bertambahnya kasus Covid-19 di Kota Surabaya. (Apriliyanti et.all, 2021)

4.8. Integrasi Keislaman

Ditengah pandemi wabah Covid-19 saat ini, untuk menjaga kesehatan dan keselamatan diri dan keluarga kita sangat penting. Dengan mematuhi protokol ke-

- Bunkers, M., Jr, J. M. DeGaetano, A. (1996). Definition of Climate Regions in the Northern Plains Using an Objective Cluster Modification Technique. *Journal of Climate*, 0442(Januari). Retrieved from [http://journals.ametsoc.org/doi/abs/10.1175/1520-0442\(1996\)009](http://journals.ametsoc.org/doi/abs/10.1175/1520-0442(1996)009)
- Chitra, K., dan Maheswari, D. (2017). A comparative study of various clustering algorithms in data mining. *International Journal of Computer Science and Mobile Computing*, 6(8), 109–115.
- E. Yulian.(2018).”Text Mining dengan K-Means Clustering pada Tema LGBT dalam Arsip Tweet Masyarakat Kota Bandung,” *Jurnal Matematika ”Mantik”*, pp. 53-58, 2018.
- Fadlina, A. (2015). Penerapan Metode Agglomerative Hierarchical Clustering Untuk Klasifikasi Kabupaten/ Kota di Provinsi Jawa Timur Berdasarkan Kualitas Pelayanan Keluarga Berencana. Malang: UIN Maulana Malik Ibrahim.
- Febriyanti, F., Hafiyusholeh, M., Asyhar, A. H. (2016). Perbandingan Pengklusteran Data Iris Menggunakan Metode K-Means dan Fuzzy C-Means . *Jurnal Matematika”MANTIK”*.
- Gujarati, D.N. dan D.C. Porter. 2010. *Dasar-Dasar Ekonometrika*, Edisi 5. Jakarta: Salemba Empat
- Han, J., Kamber, M., Pei, J. (2012). *Data Mining Concept and Techniques*, 3rd. Amsterdam: Morgan Kaufmann-Elsevier.
- Handoyo, Rendy, R Rumani, and Surya Michrandi Nasution. (2014). “Perbandingan Metode Clustering Menggunakan Metode Single Linkage Dan K-Means Pada Pengelompokan Dokumen.” *Jsm Stmik Mikroskil* 15(2): 73–82.

- Rahmawati, L., Widya Sihwi, S., Suryani, E. (2016). Analisa Clustering Menggunakan Metode K-Means Dan Hierarchical Clustering (Studi Kasus : Dokumen Skripsi Jurusan Kimia, Fmipa, Universitas Sebelas Maret). *Jurnal Teknologi Informasi ITSmart*, 3(2), 66. <https://doi.org/10.20961/its.v3i2.654>
- Raftery, A. E., dan Dean, N. (2006). Variable selection for model-based clustering. *Journal of the American Statistical Association*, 101(473), 168–178.
- Rani, Y., dan Rohil, H. (2013). A study of hierarchical clustering algorithm. *International Journal of Information and Computation Technology*, 3(11), 1225–1232.
- Rousseeuw, P. J. (1984). Least Median of Squares Regression. *Journal of the American Statistical Association*, 871-880.
- Satria, F., Aziz A. RZ. (2016). Perbandingan Kinerja Metode Ward dan K-Means dalam Menentukan Cluster Data Mahasiswa Permohonan Beasiswa (STUDI KASUS : STMIK PRINGSEWU). *Jurnal TIM Darmajaya*, 2, 1.
- Santoso, Slamet, I. (1987). *Pendidikan di Indonesia dari Masa ke Masa*. Jakarta: CV. Haji Masagung.
- Silvi, R. (2018). Analisis Cluster dengan Data Outlier Menggunakan Centroid Linkage dan K-Means Clustering untuk Pengelompokkan Indikator HIV/AIDS di Indonesia. *Jurnal Matematika "MANTIK,"* 4(1), 22–31. <https://doi.org/10.15642/mantik.2018.4.1.22-31>
- Sitompul, B. D. (2018). Peningkatan Hasil Evaluasi Clustering Davies-Bouldin Index dengan Penentuan Titik Cluster Awal Algoritma K-Means. Medan: Universitas Sumatra Utara.

