

PENGELOMPOKAN KABUPATEN/KOTA DI JAWA TIMUR BERDASARKAN INDIKATOR KASUS DBD MENGGUNAKAN *COMPLETE LINKAGE* DAN *AVERAGE LINKAGE*

Mutia Eva Mustafidah*

Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya

Mohammad Dian Purnama

Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya

*Corresponding author: mutia.20010@mhs.unesa.ac.id

Abstrak

Demam Berdarah Dengue (DBD) merupakan permasalahan serius dalam sektor kesehatan masyarakat Indonesia, dengan tingkat kematian yang cenderung meningkat setiap tahunnya. Menurut data, kasus DBD tersebar di 449 kabupaten/kota atau sekitar 87,3% dari total wilayah administratif di 34 provinsi Indonesia. Pada tahun 2022, Dinas Kesehatan mencatat tingginya jumlah kasus DBD, mencapai 131.265 kasus. Provinsi Jawa Timur menduduki peringkat kedua dalam jumlah kasus DBD tertinggi, dengan 8.894 kasus tercatat di wilayah tersebut. Oleh karena itu, penelitian ini dilakukan untuk mengidentifikasi faktor-faktor yang berpengaruh terhadap kasus DBD menggunakan sembilan indikator, termasuk persentase penduduk miskin, kepadatan penduduk, persentase Indeks Pembangunan Manusia (IPM), jumlah sarana kesehatan, persentase keluhan kesehatan, persentase rumah tangga dengan sanitasi layak, persentase tempat umum yang memenuhi syarat kesehatan, persentase air minum layak, dan persentase jaminan kesehatan. Algoritma yang digunakan dalam penelitian ini adalah algoritma hierarki clustering dengan metode *complete linkage* dan *average linkage*. Tujuan utama dari penelitian ini adalah untuk menghasilkan kelompok-kelompok Kabupaten/Kota di Provinsi Jawa Timur berdasarkan indikator yang berkaitan dengan kasus DBD. Hasil penelitian menunjukkan adanya empat kelompok, yakni kelompok 1 dengan 28 Kabupaten/Kota, kelompok 2 dengan 5 Kabupaten/Kota, kelompok 3 dengan 4 Kabupaten/Kota, dan kelompok 4 dengan 1 Kabupaten/Kota.

Kata Kunci: *Clustering, complete linkage, average linkage, Demam Berdarah Dengue.*

Abstract

Dengue Hemorrhagic Fever (DHF) is a serious problem in Indonesia's public health sector, with a mortality rate that tends to increase every year. According to data, DHF cases are spread across 449 districts/cities or around 87.3% of the total administrative area in 34 provinces of Indonesia. In 2022, the Health Office recorded a high number of dengue cases, reaching 131,265 cases. East Java Province ranked second in the highest number of DHF cases, with 8,894 cases recorded in the region. Therefore, this study was conducted to identify factors that influence dengue cases using nine indicators, including the percentage of poor people, population density, percentage of Human Development Index (HDI), number of health facilities, percentage of health complaints, percentage of households with proper sanitation, percentage of public places that meet health requirements, percentage of proper drinking water, and percentage of health insurance. The algorithm used in this research is a hierarchical clustering algorithm with *complete linkage* and *average linkage* methods. The main objective of this study was to generate groups of districts/cities in East Java Province based on indicators related to dengue cases. The results showed four groups, namely group 1 with 28 districts/cities, group 2 with 5 districts/cities, group 3 with 4 districts/cities, and group 4 with 1 district/city.

Keywords: *Clustering, complete linkage, average linkage, Dengue Hemorrhagic Fever.*

PENDAHULUAN

Demam Berdarah Dengue (DBD) adalah penyakit yang disebabkan oleh virus dengue yang ditransmisikan melalui gigitan nyamuk *Aedes Aegypti* (Mubarak & Budiantara, 2012). Virus dengue menyerang sistem kekebalan tubuh, sehingga dapat menurunkan daya tahan tubuh

dalam waktu singkat. DBD telah lama menjadi penyakit menular yang menjadi salah satu penyebab utama kematian di seluruh dunia (Hendriana, 2022). Pada tahun 2021, Organisasi Kesehatan Dunia (WHO) memperkirakan terjadi sekitar 100 hingga 400 juta infeksi DBD setiap tahun secara global. DBD merupakan permasalahan serius dalam kesehatan masyarakat Indonesia, dan angka kematian

akibatnya cenderung meningkat dari tahun ke tahun. Di Indonesia, kasus DBD menyebar di 449 kabupaten/kota atau sekitar 87,3% dari total kabupaten/kota di 34 provinsi. Upaya pencegahan penyakit DBD merupakan langkah terbaik yang dapat diambil melalui kegiatan preventif.

Berdasarkan data Dinas Kesehatan Jawa Timur, jumlah kasus Demam Berdarah Dengue (DBD) di Indonesia masih tinggi, mencapai 131.265 kasus di tahun 2022. Provinsi Jawa Timur menempati peringkat kedua dengan jumlah kasus DBD tertinggi di seluruh Indonesia. DBD menjadi permasalahan kesehatan masyarakat yang merata di hampir semua kabupaten atau kota di Provinsi Jawa Timur. Informasi dari Dinas Kesehatan Provinsi Jawa Timur menunjukkan bahwa jumlah kasus DBD mengalami penurunan dari 8.567 kasus pada tahun 2020 menjadi 6.760 kasus pada tahun 2021. Akan tetapi, terjadi peningkatan yang signifikan pada tahun 2022, dengan penambahan sebanyak 2.134 kasus, sehingga tercatat total 8.894 kasus DBD di wilayah Jawa Timur.

Metode pengelompokan yang sering digunakan dalam penelitian adalah *clustering analysis* (Sibuea & Sapta, 2017). *Clustering analysis* merupakan salah satu teknik pengelompokan untuk analisis pada data mining (Paramadina, 2018). Proses pengelompokan satu objek fisik atau abstrak ke dalam kelas-kelas objek serupa dikenal sebagai clustering. Teknik pengelompokan berperan penting dalam machine learning, data mining, pengambilan informasi, pemasaran, diagnostik medis, dan pengenalan pola. *Clustering* sering disebut tugas belajar tanpa pengawasan karena tidak ada kelas yang menunjukkan nilai dari clustering sebelumnya yang diberikan dari sampel data (Paramadina, 2018). Algoritma yang akan digunakan pada penelitian ini adalah *algoritma hierarchial clustering* dengan metode pengelompokan yang digunakan adalah *complete linkage* dan *average linkage* untuk mengklaster atau mengelompokkan data (Sabara, 2022).

Menurut penelitian sebelumnya persentase penduduk miskin, kepadatan penduduk, persentase IPM, persentase sarana kesehatan, persentase keluhan kesehatan merupakan faktor yang berpengaruh terhadap kasus DBD (Mubarak & Budiantara, 2012). Penelitian lainnya persentase rumah tangga dengan sanitasi layak, persentase tempat-tempat umum yang memenuhi syarat

kesehatan, persentase air minum layak merupakan faktor-faktor yang berpengaruh signifikan terhadap kasus DBD (Hendriana, 2022). Selain itu, menurut Dinas Kesehatan persentase jaminan kesehatan merupakan salah satu faktor yang berpengaruh terhadap kasus DBD. Berdasarkan uraian di atas, penelitian ini mengelompokkan Kabupaten/Kota di Jawa Timur berdasarkan indikator-indikator yang berpengaruh terhadap kasus DBD menggunakan *algoritma hierarchial clustering* dengan metode pengelompokan yang digunakan adalah *complete linkage* dan *average linkage* yang menggunakan sembilan indikator yaitu persentase penduduk miskin (X_1), persentase rumah tangga dengan sanitasi layak (X_2), persentase jaminan kesehatan (X_3), persentase tempat-tempat umum yang memenuhi syarat kesehatan (X_4), persentase air minum layak (X_5), kepadatan penduduk (X_6), jumlah sarana Kesehatan (X_7), persentase IPM (X_8), persentase keluhan kesehatan (X_9).

KAJIAN TEORI

MULTIKOLINEARITAS

Uji multikolinieritas digunakan untuk mengidentifikasi adanya hubungan atau korelasi antara variabel-variabel independen (bebas) dalam analisis regresi. Pengujian ini melibatkan pengamatan terhadap Variance Inflation Factor (VIF) dari setiap variabel independen terhadap variabel dependen (Gujarati, 2003). Apabila nilai VIF dari setiap variabel bebas < 10 , maka model regresi dianggap bebas dari gejala multikolinieritas. Sebaliknya, jika nilai VIF menunjukkan angka yang lebih tinggi, maka hal ini mengindikasikan adanya gejala multikolinieritas dalam model regresi tersebut (Anton, 1993). Jika koefisien korelasi antara masing-masing variabel bebas lebih dari 0,8 maka terjadi multikolinearitas dan sebaliknya, jika koefisien korelasi antara masing-masing variabel bebas kurang dari 0,8 maka tidak terjadi multikolinearitas (Gujarati, 2003). Hipotesis yang digunakan dalam uji multikolinearitas yaitu :

H_0 = Tidak terdapat multikolinearitas

H_1 = Terdapat multikolinearitas

Melalui pengujian kriteria sebagai berikut :

Jika nilai koefisien korelasi $> 0,8$ maka H_0 ditolak, artinya terdapat multikolinearitas. Jika nilai koefisien korelasi $< 0,8$ maka H_0 diterima, artinya tidak terdapat multikolinearitas.

EUCLIDEAN DISTANCE

Dalam pengelompokan data, perhitungan jarak antara titik-titik data dapat menggunakan *Euclidean Distance*. *Euclidean Distance* merupakan metode penghitungan jarak yang digunakan untuk mengukur jarak antara dua titik dalam ruang *Euclidean* (termasuk ruang *Euclidean* dua dimensi, tiga dimensi, atau bahkan lebih) (Anton, 1993). Untuk mengukur tingkat kesamaan data, rumus *Euclidean distance* digunakan, dan rumus ini dihitung seperti berikut:

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$$

Dimana:

d_{ij} = Jarak *Euclidean* dari objek ke-i dan j
 n = jumlah variabel
 x_{ik} = Nilai objek ke-i pada variable ke-k
 x_{jk} = Nilai objek ke-j pada variable ke-k

CLUSTERING

Clustering merupakan salah satu teknik dari salah satu fungsionalitas data mining, algoritma clustering merupakan algoritma pengelompokan sejumlah data menjadi kelompok-kelompok data tertentu (cluster) (Sibuea & Sapta, 2017). Pada proses clustering, tahapan menentukan atau mendeskripsikan nilai kuantitatif dari tingkat kemiripan atau ketidakmiripan data (proximity measure) memiliki peranan sangat penting, sehingga perlu dilakukannya perbandingan beberapa metode yang sering digunakan, yaitu jarak euclidean, manhattan, dan minkowski (Nishom, 2017). Clustering adalah peroses pengelompokan objek ke dalam class dengan mengetahui kesamaan objek. Sebuah cluster adalah kumpulan data yang mempunyai persamaan satu dengan yang lain, dalam cluster yang mempunyai kesamaan dan memiliki perbedaan dengan cluster yang lain. Clustering merupakan pembelajaran yang tidak dapat ditebak karena tidak dibutuhkan target output atau yang sering disebut unsupervised learning. Clustering dibagi menjadi 2 metode yaitu: hierarchical dan non-hierarchical (Paramadina, 2018). Hierarchical adalah metode pengelompokan data yang diawali dengan mengelompokkan beberapa objek yang memiliki kesamaan paling dekat, kemudian dilanjutkan kepada objek lain yang memiliki jarak yang lebih dekat, sehingga cluster

akan membentuk hirarki dari yang paling mirip dan yang tidak mirip (Muttaqin, 2023).

HIERARCHICAL AGGLOMERATIVE CLUSTERING

Salah satu algoritma pada *clustering hierarchical* adalah HAC (*Hierarchical Agglomerative Clustering*) yaitu metode pengelompokan data yang menggabungkan dua buah klaster yang mempunyai kemiripan (Muttaqin, 2019). *Agglomerative hierarchical clustering* (AHC) dengan menggunakan bottom-up, dimulai dari masing-masing data sebagai sebuah cluster, kemudian secara rekrusif mencar kelompok terdekat sebagai pasangan yang kemudian akan digabungkan menjadi kelompok yang lebih besar. Proses tersebut diulang terus sehingga tampak bergerak keatas membentuk hirarki (Paramadina, 2018). Langkah Algoritma Agglomerative Hierarchical Clustering sebagai berikut (Septianingsih, 2022):

1. Menghitung matriks jarak antar data menggunakan Euclidian distance.
2. Menggabungkan dua cluster dengan jarak paling dekat berdasarkan parameter kedekatan yang ditentukan
3. Memperbarui matriks jarak antar data untuk merepresentasikan kedekatan diantara kelompok baru yang terbentuk dari gabungan dua cluster dan kelompok yang masih tersisa
4. Ulangi Langkah ke-3 dan ke-4 hingga hanya satu clatter yang tersisa

Metode *Agglomeratif Hierarchical Clustering* yang digunakan pada penelitian ini adalah *complete linkage* dan *average linkage*.

Complete Linkage

Metode *complete linkage* adalah metode yang dilakukan dengan mengelompokkan data berdasarkan jarak antara cluster A dan B didefinisikan sebagai jarak maksimum antara titik di A dan titik di B (Wahyuni & Jatmiko, 2018). Metode *complete linkage* dapat dihitung dengan persamaan berikut:

$$D(A,B) = \max\{d(y_i, y_j)\}, A \in y_i, B \in y_j$$

Dimana $d(y_i, y_j)$ merupakan jarak *Euclidean* atau jarak lain antara vektor y_i dan y_j . Pada setiap langkah metode *complete linkage*, jarak $D(A,B)$ ditentukan untuk setiap pasangan cluster dan dua cluster dengan jarak terbesar yang digabungkan (Rencher, 2002).

Average Linkage

Metode *average linkage* merupakan metode yang dilakukan dengan mengelompokkan data

berdasarkan jarak rata-rata antar keseluruhan data. Dalam pendekatan *average linkage*, jarak antara dua kluster A dan B didefinisikan sebagai rata-rata jarak $n_A n_B$ antara titik n_A di A dan titik n_B di B (Suhirman & Wintolo, 2019). Metode *average linkage* dapat dihitung dengan persamaan berikut

$$D(A, B) = \frac{1}{n_A n_B} \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} d(y_i, y_j)$$

Dimana jumlah total y_i di A dan y_j di B. Pada setiap langkah *average linkage*, dengan menggunakan $D(A, B)$ untuk mencari gabungan kedua *cluster* dengan jarak terkecil (Rencher, 2002).

METODE

Penelitian ini menggunakan data sekunder yang berasal dari BPS Jawa Timur. Variabel pada penelitian ini meliputi persentase penduduk miskin (X_1), persentase rumah tangga dengan sanitasi layak (X_2), persentase jaminan kesehatan (X_3), persentase tempat-tempat umum yang memenuhi syarat kesehatan (X_4), persentase air minum layak (X_5), kepadatan penduduk (X_6), jumlah sarana Kesehatan (X_7), persentase IPM (X_8), persentase keluhan kesehatan (X_9). Sembilan indikator tersebut merupakan variabel yang digunakan dalam proses pengelompokkan pada 38 kabupaten/kota di Jawa Timur. Adapun tahap analisis pada penelitian ini sebagai berikut:

1. Melakukan pengumpulan data yang diambil melalui laman BPS Provinsi Jawa Timur 2022
2. Melakukan uji multikolineritas
3. Melakukan pengelompokkan dengan metode *complete linkage*
4. Melakukan pengelompokkan dengan metode *average linkage*
5. Menginterpretasikan hasil *cluster* yang terbentuk.

HASIL DAN PEMBAHASAN

UJI MULTIKOLINEARITAS

Data pada penelitian ini terdapat 9 variabel sehingga output matriks yang dihasilkan adalah matriks korelasi 9×9 pada tabel berikut:

Tabel 1. Hasil Uji Multikolineritas 1

	X_1	X_2	X_3	X_4	X_5
X_1	1,00	-0,61	-0,55	-0,57	-0,51
X_2	-0,61	1,00	0,57	0,46	0,36
X_3	-0,55	0,57	1,00	0,37	0,36
X_4	-0,57	0,46	0,37	1,00	0,49
X_5	-0,51	0,36	0,36	0,49	1,00
X_6	-0,56	0,48	0,46	0,45	0,43
X_7	-0,03	0,01	-0,12	-0,26	-0,04
X_8	-0,81	0,75	0,62	0,61	0,44
X_9	-0,01	0,03	-0,07	-0,03	-0,27

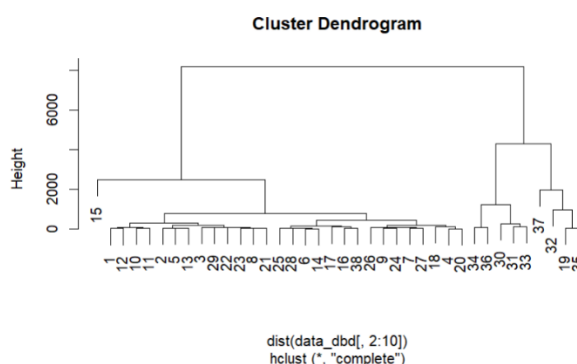
Tabel 2. Hasil Uji Multikolineritas 2

	X_6	X_7	X_8	X_9
X_1	-0,55	-0,03	-0,81	-0,01
X_2	0,48	0,01	0,75	0,02
X_3	0,46	-0,12	0,62	-0,07
X_4	0,45	-0,26	0,61	-0,03
X_5	0,43	-0,04	0,44	-0,27
X_6	1,00	-0,03	-0,81	-0,01
X_7	-0,05	1,00	0,02	-0,18
X_8	0,72	0,017	1,00	-0,23
X_9	-0,27	-0,18	-0,23	1,00

Pada tabel di atas menunjukkan nilai korelasi antar variabel semuanya kurang dari 0,8 sehingga dapat disimpulkan bahwa tidak terjadi multikolineritas dan semua variabel tersebut dapat digunakan untuk analisis *cluster*.

COMPLETE LINKAGE

Setelah melakukan proses klasterisasi menggunakan *complete linkage* yang hasilnya dapat dijelaskan dengan menggambarannya dalam bentuk dendrogram. Hasil klasterisasi metode *complete linkage* ditunjukkan pada gambar berikut :



Gambar 1. Cluster dendrogram metode *complete linkage*

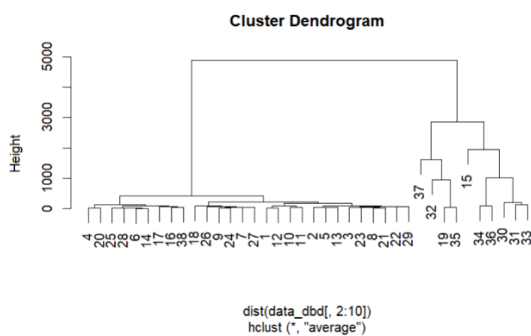
Berdasarkan dendrogram metode *complete linkage* diatas maka apabila dibentuk 4 *cluster* didapatkan hasil sebagai berikut:

Tabel 3. Hasil *Clustering* metode *complete linkage*

	Anggota	Jumlah
Cluster 1	1,2,3,4,5,6,7,8,9,10,11, 12,13,14,16,17,18,20,21, 22,23,24,25,26,27,28,29,38	28
Cluster 2	30,31,33,34,36	5
Cluster 3	19,32,35,37	4
Cluster 4	15	1

AVERAGE LINKAGE

Setelah melakukan proses klasterisasi menggunakan *average linkage* yang hasilnya dapat dijelaskan dengan menggambarannya dalam bentuk dendrogram sebagai berikut :



Gambar 2. Cluster dendrogram metode *average linkage*

Berdasarkan dendrogram metode *average linkage* diatas maka 4 *cluster* yang didapatkan sebagai berikut:

Tabel 4. Hasil *Clustering* metode *complete linkage*

	Anggota	Jumlah
Cluster 1	1,2,3,4,5,6,7,8,9,10,11, 12,13,14,16,17,18,20,21, 22,23,24,25,26,27,28,29,38	28
Cluster 2	30,31,33,34,36	5
Cluster 3	19,32,35,37	4
Cluster 4	15	1

KARAKTERISTIK ANGGOTA CLUSTER

Dari kedua metode dengan empat *cluster*, didapatkan anggota hasil *clustering* sama dan tidak ada perbedaan sama sekali. Setelah didapatkan anggota masing-masing *cluster*, langkah selanjutnya adalah perhitungan rata-rata karakteristik variable atau indikator pada masing-masing *cluster*. Dengan menghitung rata-rata, peneliti atau analis data memiliki kesempatan untuk menjelajahi lebih mendalam tentang apa yang membuat setiap *cluster* berbeda satu sama lain. Hal ini memberikan tentang

karakteristik unik yang membedakan satu kelompok dari yang lain.

Tabel 5. Hasil rata-rata pada masing-masing indikator 1

	X_1	X_2	X_3	X_4	X_5
Cluster 1	11,74	78,67	62,57	84,67	66,35
Cluster 2	6,48	93,15	95,47	94,55	87,98
Cluster 3	6,47	91,00	78,80	93,36	86,12
Cluster 4	5,36	89,54	77,65	83,7	43,69

Tabel 6. Hasil rata-rata pada masing-masing indikator 2

	X_6	X_7	X_8	X_9
Cluster 1	785,86	36,29	70,89	8,51
Cluster 2	4853,6	11,20	78,53	35,26
Cluster 3	7381,2	43,25	79,29	24,91
Cluster 4	2905	51	81,02	21,92

Berdasarkan tabel diatas diperoleh interpretasi karakteristik pada kedua metode memperoleh hasil anggota *cluster* yang sama, sehingga interpretasi untuk setiap *cluster* adalah sebagai berikut :

1. Cluster 1 beranggotakan 28 kabupaten/kota, yaitu: Kabupaten Kediri, Kabupaten Banyuwangi, Kabupaten Pasuruan, Kabupaten Magetan, Kabupaten Pacitan, Kabupaten Trenggalek, Kabupaten Situbondo, Kabupaten Bondowoso, Kabupaten Tulungagung, Kabupaten Lumajang, Kabupaten Sampang, Kabupaten Nganjuk, Kabupaten Ngawi, Kabupaten Ponorogo, Kabupaten Bangkalan, Kabupaten Pamekasan, Kabupaten Mojokerto, Kabupaten Jombang, Kabupaten Lamongan, Kabupaten Blitar, Kabupaten Probolinggo, Kabupaten Tuban, Kabupaten Sumenep, Kabupaten Bojonegoro, Kabupaten Gresik, Kabupaten Malang, Kabupaten Jember, Kota Batu. Pada *cluster* 1 persentase penduduk miskin paling tinggi dibandingkan cluster yang lain yaitu 11,74. Untuk persentase sanitasi layak sebesar 78,67, persentase jaminan kesehatan sebesar 62,57 dan persentase IPM sebesar 70,89, memiliki nilai paling rendah dibandingkan *cluster* lainnya sehingga diperlukan perhatian khusus karena semakin rendah ketiga indikator tersebut maka semakin tinggi kasus DBD. Sedangkan persentase tempat-tempat umum yang memenuhi syarat Kesehatan sebesar 84,67, persentase air minum layak sebesar 66,35, jumlah sarana kesehatan sebesar 36,29 sehingga ketiga indikator

tersebut tergolong cukup rendah karena menempati urutan kedua terendah dibandingkan *cluster* lainnya yang menyebabkan kasus DBD semakin tinggi di wilayah tersebut. Kemudian untuk kepadatan penduduk sebesar 785,86 dan persentase keluhan Kesehatan sebesar 8,51.

2. *Cluster* 2 beranggotakan 5 kabupaten/kota, yaitu: Kota Kediri, Kota Blitar, Kota Probolinggo, Kota Pasuruan, Kota Madiun. Pada *cluster* 2 persentase sanitasi layak sebesar 93,15, persentase jaminan Kesehatan sebesar 95,47, persentase tempat-tempat umum yang memenuhi syarat kesehatan sebesar 94,55, persentase air minum layak sebesar 87,98, keempat indikator tersebut paling tinggi dibandingkan *cluster* lainnya. Sedangkan persentase penduduk miskin sebesar 6,48 dan kepadatan penduduk sebesar 4853,60 menempati urutan kedua tertinggi dibandingkan *cluster* lainnya sehingga diperlukan perhatian khusus karena semakin tinggi kedua indikator tersebut maka semakin tinggi kasus DBD. Jumlah sarana kesehatan paling rendah dibandingkan *cluster* lainnya dan persentase IPM sebesar 78,53 juga cukup tinggi. Kemudian untuk persentase keluhan kesehatan sebesar 35,26 paling tinggi dibandingkan *cluster* lainnya sehingga diperlukan perhatian khusus karena keluhan Kesehatan akan berpengaruh terhadap kenaikan kasus DBD.
3. *Cluster* 3 beranggotakan 4 kabupaten/kota, yaitu: Kabupaten Madiun, Kota Malang, Kota Mojokerto, Kota Surabaya. Pada *cluster* 3 persentase penduduk miskin sebesar 6,47, persentase sanitasi layak sebesar 91,00, persentase jaminan Kesehatan sebesar 78,80, persentase IPM sebesar 79,29, persentase tempat-tempat umum yang memenuhi syarat kesehatan sebesar 93,36, persentase air minum layak sebesar 86,12 dan jumlah sarana kesehatan sebesar 43,25. Sedangkan kepadatan penduduk sebesar 7381,25 paling tinggi dibandingkan *cluster* lainnya dan persentase keluhan kesehatan sebesar 24,91 berada di urutan kedua tertinggi dibandingkan *cluster* lainnya, sehingga memerlukan perhatian khusus karena semakin tinggi kedua indikator tersebut maka semakin tinggi kasus DBD.
4. *Cluster* 4 beranggotakan 1 Kabupaten yaitu : Kabupaten Sidoarjo. Pada *cluster* 4 persentase tempat-tempat umum yang memenuhi syarat kesehatan sebesar 83,7 dan persentase air minum layak sebesar 46,69 paling rendah dibandingkan *cluster* lainnya sehingga

memiliki resiko yang tinggi terhadap kasus DBD. Selain itu, persentase sanitasi layak sebesar 89,54 dan persentase jaminan kesehatan sebesar 77,65 juga perlu diperhatikan karena menempati urutan kedua terendah dibandingkan *cluster* lainnya sehingga memerlukan perhatian khusus juga untuk mencegah kenaikan kasus DBD. Kemudian persentase penduduk miskin sebesar 5,36, kepadatan penduduk sebesar 2905, jumlah sarana kesehatan sebesar 51, persentase IPM sebesar 81,02 dan persentase keluhan kesehatan sebesar 21,92.

PENUTUP

SIMPULAN

Berdasarkan hasil dan pembahasan yang telah diuraikan, maka dapat disimpulkan bahwa hasil empat kelompok atau *cluster* yang diperoleh dari dua metode yang berbeda menghasilkan anggota *cluster* yang sama. Adapun hasil analisis *cluster* yang dilakukan tersebut adalah sebagai berikut:

1. *Cluster* 1 terdiri dari 28 Kabupaten/Kota yang mana pada *cluster* ini indikator yang perlu diperhatikan pada kasus DBD adalah persentase penduduk miskin, persentase sanitasi layak, persentase jaminan Kesehatan, persentase IPM, persentase tempat-tempat umum yang memenuhi syarat Kesehatan, persentase air minum layak dan jumlah sarana Kesehatan.
2. *Cluster* 2 yang terdiri dari 5 Kabupaten/Kota yang mana pada *cluster* ini indikator yang perlu diperhatikan pada kasus DBD adalah jumlah sarana kesehatan, persentase keluhan Kesehatan, persentase penduduk miskin, kepadatan penduduk, persentase IPM.
3. *Cluster* 3 yang terdiri dari 4 Kabupaten/Kota yang mana pada *cluster* ini indikator yang perlu diperhatikan pada kasus DBD adalah kepadatan penduduk dan persentase keluhan Kesehatan.
4. *Cluster* 4 yang terdiri dari 1 Kabupaten/Kota yang mana pada *cluster* ini indikator yang perlu diperhatikan pada kasus DBD adalah persentase sanitasi layak, persentase jaminan Kesehatan, persentase tempat-tempat umum yang memenuhi syarat Kesehatan, persentase air minum layak.

SARAN

Dari hasil studi yang telah dilakukan, disarankan untuk menjelajahi variabel tambahan yang dapat

mendukung indikator kasus Demam Berdarah Dengue (DBD) dalam penelitian berikutnya. Selain itu, dipertimbangkan untuk memilih topik lain yang memiliki signifikansi dan kontroversi agar penelitian dapat memberikan kontribusi yang substansial.

DAFTAR PUSTAKA

- Dinas Kesehatan. (2022). Profil Kesehatan Jawa Timur 2022 Jawa Timur: Dinas Kesehatan
- Sibuea, M. L., & Safta, A. (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, 4(1), 85-92.
- Gujarati, D. N. (2003). *Basic Econometrics*. Third Edition. Mc. . Graw-Hill: Singapore.
- Anton, H. (1993). *Elementary Linear Algebra*, 7th ed. New Jersey: Wiley.
- Hendriana, L., (2022). Penanganan Overdispersi pada Kasus Jumlah Kematian Akibat Demam Berdarah Dengue di Provinsi Jawa Timur.
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika*, 4(01), 20-24.
- Mubarak, R., & Budiantara, I. N. (2012). Analisis Regresi Spline Multivariabel untuk Pemodelan Kematian Penderita Demam Berdarah Dengue (DBD) di Jawa Timur. *Jurnal Sains dan Seni ITS*, 1(1), D224-D229.
- Muttaqin, Z. (2023). Implementasi Unsupervised Learning Pada Nilai Jasmani Kesamaptaan Sekolah Polisi Negara Dengan Metode Clustering Analysis. *PROSISKO: Jurnal Pengembangan Riset dan Observasi Sistem Komputer*, 10(1), 18-23.
- Paramadina, M. (2018). Perbandingan Analisis Cluster Metode Average Linkage dan Metode Ward (Kasus IPM Provinsi Sulawesi Selatan) (Doctoral dissertation, FMIPA).
- Rencher, A.R. (2002). *Methods of Multivariate Analysis* Second Edition. John Wiley & Sons, Inc. New York.
- Respati, T., Raksanegara, A., Djuhaeni, H., Sofyan, A., Agustian, D., Faridah, L., & Sukandar, H. (2017). Berbagai faktor yang memengaruhi kejadian demam berdarah dengue di Kota Bandung. *ASPIRATOR-Journal of Vector-borne Disease Studies*, 9(2), 91-96.
- Sabara, I. M. (2022). Analisis agglomerative hierarchical clustering berdasarkan pengurutan parsial graf hasse terhadap indikator kemiskinan di Jawa Timur (Doctoral dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).
- Septianingsih, A. (2022). PEMETAAN KABUPATEN KOTA DI PROVINSI JAWA TIMUR BERDASARKAN TINGKAT KASUS PENYAKIT MENGGUNAKAN PENDEKATAN AGGLOMERATIF HIERARCHICAL CLUSTERING. *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 3(2), 367-386.
- Sucipto, P.T., Raharjo, M. and Nurjazuli, N., (2015). Faktor-faktor yang mempengaruhi kejadian penyakit Demam Berdarah Dengue (DBD) dan jenis serotipe virus Dengue Di Kabupaten Semarang. *Jurnal Kesehatan Lingkungan Indonesia*, 14(2), pp.51-56.
- Suhirman, H.W. and Wintolo, H. (2019). System for Determining Public Health Level Using the Agglomerative Hierarchical Clustering Method. *Compiler*, 8(1), pp.95-104.
- Wahyuni, S. and Jatmiko, Y.A., (2018). Pengelompokan Kabupaten/Kota di Pulau Jawa Berdasarkan Faktor-Faktor Kemiskinan dengan Pendekatan Average Linkage Hierarchical Clustering. *Jurnal Aplikasi Statistika & Komputasi Statistik*, 10(1), pp.1-8.