

Agglomerative Nesting Cluster Analyst in Mapping District/City Health Facilities in West Java Province

Analisis *Cluster Agglomeratif Nesting* dalam Pemetaan Sarana Kesehatan Kabupaten/kota di Provinsi Jawa Barat

Megawati¹, Nadira Nisa Alwani², Anwar Fitrianto³, Erfiani⁴, Alfa Nugraha Pradana⁵

^{1,2,3,4,5}*Pogram Studi Statistika dan Sains Data, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor, Indonesia*

Email: ¹mw.ramly@gmail.com, ²nisaalwani1@gmail.com, ³anwarstat@gmail.com, ⁴erfiani@apps.ipb.ac.id, ⁵alfanugraha@apps.ipb.ac.id

Received: 27 November 2023, revised: 19 January 2024, accepted: 19 January 2024

Abstract

The use of Hierarchical Clustering is used to group districts or cities in West Java according to the number of health facilities, distance to health facilities and population density using Agglomerative Nesting (AGNES). Clustering in this study utilizes complete linkage clustering. The elbow method produces two optimal clusters which are then validated with the sillhoute coefficient and Calinski-Harabasz. In this study, there are 27 variables in the form of health facilities spread across 27 regencies/cities in West Java in 2021. The results of the cluster analysis formed in this study are 18 districts/cities in cluster one and 9 districts/cities in cluster two.

Keywords: Agglomerative Nesting; Clustering; Health Facilities

Abstrak

Pemanfaatan *Hierarchical Clustering* digunakan untuk mengelompokkan kabupaten atau kota di Jawa Barat menurut jumlah sarana Kesehatan, jarak menuju sarana Kesehatan dan kepadatan penduduk menggunakan *Agglomerative Nesting (AGNES)*. *Clustering* pada penelitian ini memanfaatkan *clustering complete linkage*. Metode elbow menghasilkan dua *cluster* optimal yang kemudian divalidasi dengan *sillhoute coefficient* dan *Calinski-Harabasz*. Pada penelitian ini terdapat 27 peubah berupa sarana kesehatan yang tersebar pada 27 kabupaten/kota di Jawa Barat tahun 2021. Hasil analisis *cluster* yang terbentuk dalam penelitian ini terdapat 18 kabupaten/kota pada *cluster* satu dan 9 kabupaten/kota pada *cluster* dua.

Kata kunci: *Agglomerative Nesting; Clustering; Sarana Kesehatan*



1. PENDAHULUAN

Data mining merupakan proses yang dilakukan untuk mendapatkan sebuah informasi yang berguna dalam penyimpanan data yang berukuran besar agar dapat memberikan bantuan dalam mengambil keputusan. *Data Mining* terbagi menjadi dua kategori yaitu prediktif dan deskriptif [21]. Dalam prediktif berguna untuk memprediksi suatu nilai dari 1 atribut tertentu berdasarkan pada atribut lainnya. Sedangkan, *data mining* untuk deskriptif berguna untuk menurunkan pola pada korelasi, *clustering*, dan trend. Penggunaan *data mining* dapat digunakan pada berbagai bidang seperti perbankan, Pendidikan, riset, sampai bidang Kesehatan.

Analisis *cluster* merupakan salah satu metode yang cukup penting dalam pengolahan data statistik, sering kali digunakan untuk mengelompokkan suatu objek kedalam sebuah *cluster*. Apabila dalam satu *cluster* memiliki nilai homogenitas yang tinggi dan heterogenitas besar antar satu *cluster* dengan *cluster* lainnya maka hasil *cluster* tersebut dapat dikatakan baik. Dengan kata lain, variasi setiap anggota dalam satu *cluster* harus rendah sedangkan variasi antar setiap *cluster* harus tinggi [18].

Pada dasarnya metode *clustering* dibagi menjadi empat bagian yaitu metode hierarki, metode berbasis kepadatan, metode berbasis *grid*, dan metode partisi [3]. Terdapat dua jenis metode hierarki dalam mengelompokkan ialah *agglomerative* dan *divisive*. Metode penggabungan atau dapat disebut dengan *agglomerative* merupakan pengelompokkan yang dimulai dari setiap objek pada satu *cluster* yang terpisah kemudian membentuk *cluster* yang semakin besar, banyaknya *cluster* awal adalah banyaknya objek. *Divisive* merupakan metode pembagian yang dimulai dari semua objek kemudian dikelompokkan menjadi *cluster* tunggal yang terpisah hingga setiap objek berada didalam *cluster* yang terpisah. Beberapa algoritma yang termasuk kedalam *agglomerative* diantaranya ialah *Single Linkage*, *Complete Linkage*, dan *Average Linkage* [1].

Algoritma *agglomerative* dimulai dengan menghitung matriks jarak antara setiap objek. Semua objek membentuk cluster dan secara bertahap digabungkan menjadi setiap pasangan *cluster* terdekat dengan menggunakan metode penggabungan berdasarkan pengukuran jarak hingga semua *cluster* digabungkan menjadi satu [9]. Analisis *cluster* ini berguna untuk mengetahui pemetaan sarana Kesehatan pada suatu aspek-aspek tertentu. Dalam suatu wilayah contohnya pedesaan pasti memiliki aspek yang berbeda satu sama lain. Dalam Potensi Desa (PODES) telah menyediakan data tentang sarana Kesehatan, potensi ekonomi, sosial dan lainnya [4].

Beberapa penelitian yang mengkaji tentang analisis *cluster* adalah [7] tentang Penerapan Metode *Agglomerative Nesting* (AGNES) pada Pengelompokan Wilayah Desa atau Kelurahan di Kabupaten Kutai Kartanegara 2018. Penelitian ini menyimpulkan terdapat dua *cluster* yang terbentuk, dengan *cluster* pertama beranggotakan 228 desa/kelurahan dan *cluster* 2 beranggotakan 4 desa/kelurahan dengan kedua *cluster* memiliki nilai *icdrate* yang dihasilkan sebesar 0.5965.

Penelitian lain oleh [10] yaitu Pengelompokkan Desa atau Kelurahan di Kutai Kartanegara Menggunakan Algoritma *Divise Analysis* Tahun 2018. Penelitian ini menyimpulkan bahwa terdapat dua *cluster* optimal yang terbentuk pada pengelompokkan desa/kelurahan. *Cluster* pertama beranggotakan 230 desa/kelurahan sedangkan *cluster* kedua beranggotakan 2 desa/kelurahan dengan kedua *cluster* memiliki nilai *shillhouette coefficient* yang dihasilkan sebesar 0.744.

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana

Lebih lanjut, penelitian yang dilakukan oleh Ulinuha dkk, membahas Penerapan *Fuzzy c-Means* dalam Mengelompokkan Fasilitas Kesehatan di [13]. Penelitian ini menyimpulkan terdapat tiga *cluster* penyebaran fasilitas Kesehatan di Jawa Timur yaitu, *cluster* dengan tingkat jumlah fasilitas kesehatan tinggi terdapat 4 kabupaten/kota, *cluster* dengan tingkat sedang terdapat 26 kabupaten/kota, dan *cluster* rendah terdapat 8 kabupaten/kota. Ketiga *cluster* tersebut memiliki nilai *silhouette Coefisient* sebesar 0.695 dan dikatakan sudah baik dalam kedekatan data didalam satu *cluster*.

Adapun penelitian yang mengkaji tentang meninjau kembali pengelompokkan *agglomerative* oleh [23] Metode aglomerasi hierarki menonjol sebagai pendekatan yang sangat efektif dan populer untuk mengelompokkan data. Namun, metode-metode ini belum dibandingkan secara sistematis mengenai isu penting mengenai positif palsu saat mencari *cluster*.

Penelitian selanjutnya dilakukan oleh [12] dengan menyimpulkan bahwa pengelompokkan hierarki diakui sebagai pendekatan tanpa pengawasan yang efisien untuk menganalisis data tanpa label. Dan dalam data mining pengelompokkan hierarki ini merupakan suatu mekanisme dalam mengelompokkan data pada skala yang berbeda dengan menggunakan dendogram.

Menurut penelitian yang dilakukan oleh [19] mengatakan pengelompokkan hierarki banyak digunakan dalam data mining. Pada kriteria hubunga *single linkage* sangat kuat dikarenakan memiliki kemungkinan penanganan dari berbagai bentuk dan kepadatan, namun sensitif terhadap *noise*, terdapat dua perbaikan dalam mengatasi *noise*. Pertama, kriteria *single linkage* memperhitungkan kepadatan lokal untuk memastikan jarak antar titik-titik inti pada setiap kelompok. Kedua, algoritma hierarki melarang adanya penggabungan *cluster* yang representatif yang lebih tinggi dari ukuran minimum setelah dilakukan identifikasi.

Beberapa kajian diatas, terdapat penelitian yang secara tidak langsung membandingkan analisis *cluster* berbasis hirarki yaitu metode *Agglomerative Nesting (AGNES)* dan Algoritma *Divise Analysis*. Kedua penelitian tersebut menganalisis data yang sama dan menghasilkan jumlah *cluster cluster* yang sama pula. Oleh karena itu, dalam penelitian ini, akan diterapkan *Agglomerative Nesting (AGNES)* dengan menerapkan studi kasus seperti penelitian yang dilakukan oleh Magfiroh dkk yaitu mengelompokkan sarana Kesehatan di Jawa Barat [10].

Sarana Kesehatan berdasarkan Undang-Undang RI No.23, 1992 merupakan tempat yang digunakan untuk menyelenggarakan Upaya Kesehatan. Upaya Kesehatan dapat diwujudkan dengan berbagai cara untuk meningkatkan Kesehatan baik dilakukan oleh pemerintah maupun Masyarakat. Dalam meningkatkan Kesehatan perlu adanya Upaya Kesehatan dan pembinaan secara menyeluruh dan terpadu. Jenis sarana Kesehatan milik pemerintah menurut SNI 03-1733-2004 yang diperlukan yaitu adanya rumah sakit umum, rumah sakit bersalin, puskesmas, dan puskesmas pembantu.

Kesehatan Masyarakat sangat berarti dalam Pembangunan kesejahteraan Masyarakat Indonesia yang harus diwujudkan agar sesuai dengan cita-cita bangsa Indonesia. Memiliki pelayanan Kesehatan yang baik pada seluruh kawasan perkotaan maupun pedesaan dapat menjadi suatu tolak ukur pada bidang Kesehatan yang dapat dilihat dari angka peningkatan ekonomi Masyarakat dalam meningkatkan produktifitas Masyarakatnya [11].

Saat ini, sumber informasi mengenai kondisi Kesehatan di Jawa Barat terkhusus sarana Kesehatan masih sangat kurang sehingga perlu adanya kajian khusus yang membahas tentang hal

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana

tersebut. Tujuan dari penelitian ini adalah untuk mengelompokkan kabupaten/kota di Provinsi Jawa Barat berdasarkan sarana kesehatan menggunakan *Cluster Agglomerative Nesting* (AGNES).

2. TINJAUAN PUSTAKA

2.1 Analisis Cluster

Analisis *cluster* merupakan suatu proses data mining yang digunakan untuk mengelompokkan sebuah objek menjadi sebuah data *cluster* yang memiliki kesamaan dan karakteristik yang dapat dibedakan dari *cluster* lainnya [2]. Tidak ada asumsi mengenai jumlah kelompok atau struktur kelompok sebelumnya. Pengelompokan dilakukan berdasarkan kesamaan atau jarak (*dissimilarity*). Tujuan dari analisis cluster ialah untuk dapat mengidentifikasi kelompok objek yang memiliki kemiripan dan karakteristik klas yang dapat dipisahkan oleh kelompok lainnya. Hasil suatu kelompok dianggap baik adalah sebagai berikut [17]:

1. Homogenitas Tingkat kesamaan yang tinggi antara setiap anggota dalam suatu kelompok (dalam suatu *cluster*).
2. Heterogenitas Terdapat perbedaan yang signifikan antara kelompok pada tingkat yang satu dengan tingkat yang lain (antar *cluster*).

Jarak merupakan salah satu faktor yang sangat berpengaruh pada hasil dari kelompok yang terbentuk, maka dibutuhkan pengukuran jarak yang biasa dilakukan yaitu jarak *Euclidean*. Konsep jarak *Euclidean* yang mengukur jarak antar setiap observasi $x_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T$ dan $x_j = [x_{j1}, x_{j2}, \dots, x_{jm}]^T$ sebagai berikut:

$$d_{ij} = \sqrt{(x_i - x_j)'(x_i - x_j)} \quad (2.1)$$

$$= \sqrt{\sum_{i=1}^p (x_i - x_j)^2} \quad (2.2)$$

Dimana:

d_{ij} : jarak antar setiap observasi ke- i dan ke- j

x_i : vector objek ke- i

x_j : vector objek ke- j

2.1.1 Agglomerative Hierarchical Clustering

Ada dua jenis metode hierarki yaitu pengelompokan hierarki aglomeratif dan pengelompokan hierarki yang memecah belah (*Divisive Hierarchical Clustering*). Pengelompokan hierarki kolektif berarti bahwa setiap objek termasuk dalam kelompok yang terpisah. Kemudian gabungkan dua grup berikut untuk mendapatkan grup yang mencakup semua objek. Diperlukan tingkat ketidaksamaan antar objek individu untuk dapat menggabungkan dua kelompok pada tahap awal ketika masing-masing kelompok hanya terdiri dari satu objek [5]. Dalam hirarki *agglomerative cluster* merupakan suatu metode yang bersifat *bottom-up* yang dapat

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana

menggabungkan n buah *cluster* sehingga membentuk satu *cluster*. Teknik yang digunakan dalam mengelompokkan antar kelompok yaitu sebagai berikut:

1. *Single Linkage*

Single Linkage merupakan suatu metode analisis *cluster* hierarki dengan melakukan pengelompokkan objek berdasarkan jarak paling dekat terlebih dahulu [6].

$$d_{w(i,j)} = \min\{d_{wi}, d_{wj}\} \quad (2.3)$$

Dimana:

$d_{w(i,j)}$: jarak terpendek antar cluster

2. *Complete Linkage*

Complete Linkage dilakukan berdasarkan jarak terjauh antar setiap objek. Jarak pada antar *cluster* merupakan sebuah hasil dari perhitungan objek yang memiliki jarak terjauh. Berikut ini adalah rumus yang digunakan dalam menentukan jarak:

$$d_{w(i,j)} = \max\{d_{wi}, d_{wj}\} \quad (2.4)$$

Dimana:

$d_{w(i,j)}$: jarak terjauh antar cluster

3. *Average Linkage*

Average Linkage merupakan metode yang menggunakan rata-rata jarak selisih individu dalam suatu kelompok yang memiliki jarak seluruh individu dalam kelompok yang lain. Berikut ini rumus yang digunakan:

$$d_{w(i,j)} = \frac{\sum_q \sum_r d_{qr}}{n_{(y)}n_{(w)}} \quad (2.5)$$

Dimana:

d_{ij} : jarak antara objek i dalam cluster

$N_{(y)}$: terdapat banyaknya item dalam cluster y

$N_{(w)}$: terdapat banyaknya item dalam cluster w

3. METODOLOGI PENELITIAN

3.1 Data dan Sumber Data

Data yang digunakan dalam penelitian ini menggunakan data sekunder yaitu Potensi Desa di Provinsi Jawa Barat pada tahun 2021. Penelitian ini menggunakan 27 variabel dengan 27 observasi tersaji pada table berikut:

Tabel 3.1. Peubah yang digunakan dalam penelitian

| Peubah | Tipe Data | Satuan |
|-----------------------------|-----------|---------|
| Kepadatan Penduduk | Numerik | Jiwa/Km |
| Rumah Sakit | Numerik | Unit |
| Rumah Sakit Bersalin | Numerik | Unit |
| Puskesmas Rawat Inap | Numerik | Unit |
| Puskesmas Non Rawat Inap | Numerik | Unit |
| Puskesmas Pembantu | Numerik | Unit |
| Poliklinik Balai Pengobatan | Numerik | Unit |

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana

| | | |
|--|---------|------|
| Praktek Dokter | Numerik | Unit |
| Rumah Bersalin | Numerik | Unit |
| Praktek Bidan | Numerik | Unit |
| Poskesdes (Pos Kesehatan Desa) | Numerik | Unit |
| Polindes (Pondok Kesehatan Desa) | Numerik | Unit |
| Apotek | Numerik | Unit |
| Toko Obat | Numerik | Unit |
| Jarak Rumah Sakit | Numerik | Km |
| Jarak Rumah Sakit Bersalin | Numerik | Km |
| Jarak Puskesmas Rawat Inap | Numerik | Km |
| Jarak Puskesmas Non Rawat Inap | Numerik | Km |
| Jarak Puskesmas Pembantu | Numerik | Km |
| Jarak Poliklinik Balai Pengobatan | Numerik | Km |
| Jarak Praktek Dokter | Numerik | Km |
| Jarak Rumah Bersalin | Numerik | Km |
| Jarak Praktek Bidan | Numerik | Km |
| Jarak Poskesdes (Pos Kesehatan Desa) | Numerik | Km |
| Jarak Polindes (Pondok Kesehatan Desa) | Numerik | Km |
| Jarak Apotek | Numerik | Km |
| Jarak Toko Obat | Numerik | Km |

Data ini terdiri dari peubah numerik yang memiliki satuan jiwa/km, unit dan km.

Data mining adalah menemukan informasi tersembunyi yang menarik dalam kumpulan besar data yang disimpan di database, gudang data, atau lokasi penyimpanan data lainnya. Teknik data mining yang digunakan mempunyai tugas untuk menemukan pola-pola baru dan bermakna dalam database yang mungkin belum diketahui [21].

Data mining merupakan bagian dari proses Penemuan Pengetahuan dalam Basis Data (KDD). Aktivitas mengumpulkan dan menggunakan data historis untuk menemukan keteraturan, pola, atau hubungan dalam kumpulan data yang besar. Hasil dari data mining ini dapat digunakan untuk memperbaiki pengambilan keputusan di masa depan. Teknik yang ditemukan dalam literatur data mining meliputi pengelompokan, klasifikasi, penambangan aturan asosiasi, jaringan saraf, dan algoritma genetika [20].

4. HASIL DAN PEMBAHASAN

4.1 Eksplorasi Data

Data yang digunakan adalah data kepadatan penduduk yang bersumber dari Badan Pusat Statistik, data PODES berupa jumlah sarana kesehatan dan jarak terdekat menuju sarana kesehatan di Jawa Barat tahun 2021. Pada tahun tersebut terdapat 13 sarana kesehatan yang ada di Jawa Barat yang tersebar di 27 Kabupaten/Kota. Selanjutnya akan dilihat korelasi yang merupakan ukuran keeratan hubungan linear antar dua peubah yang ingin diketahui. Korelasi

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana

matriks pada Gambar 4.1 dan Gambar 4.2 menunjukkan korelasi antar pasangan peubah penjelas yang digunakan pada penelitian ini.

| | | | | | | | | | | | | | | |
|---------------------------------|-------|-------|-------|-------|------|-------|-------|-------|-------|-------|-------|-------|------|-------|
| Kepadatan Penduduk | 1.00 | 0.47 | 0.14 | -0.54 | 0.11 | -0.64 | 0.09 | 0.33 | 0.03 | -0.51 | -0.53 | -0.34 | 0.17 | -0.08 |
| Jumlah RS | 0.47 | 1.00 | 0.70 | -0.05 | 0.50 | -0.27 | 0.77 | 0.65 | 0.73 | 0.18 | -0.37 | -0.23 | 0.64 | 0.63 |
| Jumlah RS Bersalin | 0.14 | 0.70 | 1.00 | 0.13 | 0.49 | -0.08 | 0.84 | 0.47 | 0.86 | 0.32 | -0.32 | -0.12 | 0.53 | 0.65 |
| Jumlah Puskesmas Rawat Inap | -0.54 | -0.05 | 0.13 | 1.00 | 0.30 | 0.79 | 0.26 | 0.20 | 0.20 | 0.73 | 0.23 | 0.36 | 0.26 | 0.37 |
| Jumlah Puskesmas non Rawat Inap | 0.11 | 0.50 | 0.49 | 0.30 | 1.00 | 0.24 | 0.63 | 0.91 | 0.37 | 0.63 | 0.14 | 0.05 | 0.87 | 0.71 |
| Jumlah Puskesmas Pembantu | -0.64 | -0.27 | -0.08 | 0.79 | 0.24 | 1.00 | -0.01 | 0.01 | 0.02 | 0.66 | 0.53 | 0.41 | 0.06 | 0.13 |
| Jumlah Poliklinik | 0.09 | 0.77 | 0.84 | 0.26 | 0.63 | -0.01 | 1.00 | 0.65 | 0.87 | 0.60 | -0.23 | -0.08 | 0.78 | 0.82 |
| Jumlah Praktek Dokter | 0.33 | 0.65 | 0.47 | 0.20 | 0.91 | 0.01 | 0.65 | 1.00 | 0.37 | 0.48 | -0.09 | -0.01 | 0.92 | 0.69 |
| Jumlah Rumah Bersalin | 0.03 | 0.73 | 0.86 | 0.20 | 0.37 | 0.02 | 0.87 | 0.37 | 1.00 | 0.42 | -0.27 | -0.05 | 0.49 | 0.62 |
| Jumlah Praktek Bidan | -0.51 | 0.18 | 0.32 | 0.73 | 0.63 | 0.66 | 0.60 | 0.48 | 0.42 | 1.00 | 0.35 | 0.27 | 0.63 | 0.68 |
| Jumlah Poskesdes | -0.53 | -0.37 | -0.32 | 0.23 | 0.14 | 0.53 | -0.23 | -0.09 | -0.27 | 0.35 | 1.00 | 0.24 | 0.01 | -0.05 |
| Jumlah Polindes | -0.34 | -0.23 | -0.12 | 0.36 | 0.05 | 0.41 | -0.08 | -0.01 | -0.05 | 0.27 | 0.24 | 1.00 | 0.06 | 0.03 |
| Jumlah Apotek | 0.17 | 0.64 | 0.53 | 0.26 | 0.87 | 0.06 | 0.78 | 0.92 | 0.49 | 0.63 | 0.01 | 0.06 | 1.00 | 0.72 |
| Jumlah Toko Obat Jamu | -0.08 | -0.63 | 0.65 | 0.37 | 0.71 | 0.13 | 0.82 | 0.69 | 0.62 | 0.68 | -0.05 | 0.03 | 0.72 | 1.00 |

Gambar 4.1. Korelasi antara kepadatan Penduduk dengan jumlah sarana kesehatan

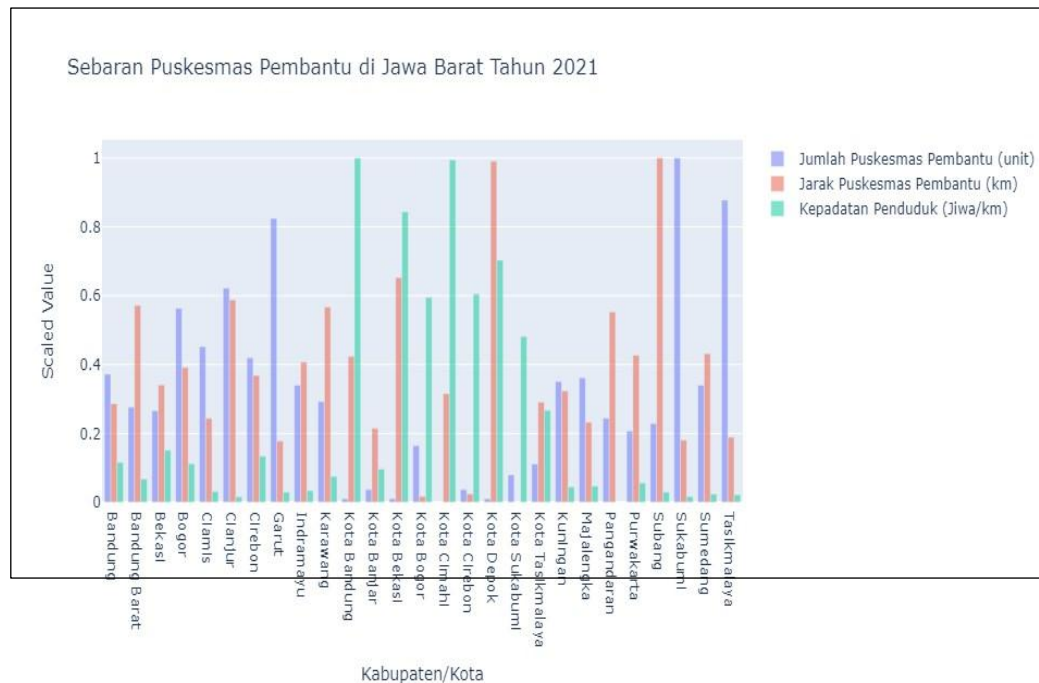
| | | | | | | | | | | | | | | |
|--------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Kepadatan Penduduk | 1.00 | -0.69 | -0.34 | -0.30 | -0.68 | 0.00 | -0.66 | -0.76 | -0.24 | -0.38 | 0.67 | 0.10 | -0.75 | -0.55 |
| Jarak RS | -0.69 | 1.00 | 0.40 | 0.34 | 0.93 | 0.07 | 0.79 | 0.92 | 0.20 | 0.58 | -0.41 | -0.04 | 0.89 | 0.76 |
| Jarak RS Bersalin | -0.34 | 0.40 | 1.00 | 0.22 | 0.38 | -0.13 | 0.52 | 0.44 | 0.74 | 0.36 | -0.34 | 0.15 | 0.46 | 0.38 |
| Jarak Puskesmas Rawat Inap | -0.30 | 0.34 | 0.22 | 1.00 | 0.33 | 0.02 | 0.53 | 0.46 | -0.04 | 0.17 | 0.12 | 0.26 | 0.48 | 0.41 |
| Jarak Puskesmas non Rawat Inap | -0.68 | 0.93 | 0.38 | 0.33 | 1.00 | 0.21 | 0.82 | 0.94 | 0.17 | 0.57 | -0.40 | -0.02 | 0.91 | 0.79 |
| Jarak Puskesmas Pembantu | 0.00 | 0.07 | -0.13 | 0.02 | 0.21 | 1.00 | -0.05 | 0.06 | -0.16 | -0.30 | 0.29 | 0.17 | 0.03 | 0.08 |
| Jarak Poliklinik | -0.66 | 0.79 | 0.52 | 0.53 | 0.82 | -0.05 | 1.00 | 0.90 | 0.30 | 0.51 | -0.43 | -0.02 | 0.92 | 0.84 |
| Jarak Praktek Dokter | -0.76 | 0.92 | 0.44 | 0.46 | 0.94 | 0.06 | 0.90 | 1.00 | 0.24 | 0.56 | -0.48 | -0.03 | 0.97 | 0.86 |
| Jarak Rumah Bersalin | -0.24 | 0.20 | 0.74 | -0.04 | 0.17 | -0.16 | 0.30 | 0.24 | 1.00 | 0.17 | -0.33 | 0.41 | 0.23 | 0.18 |
| Jarak Praktek Bidan | -0.38 | 0.58 | 0.36 | 0.17 | 0.57 | -0.30 | 0.51 | 0.56 | 0.17 | 1.00 | -0.43 | -0.04 | 0.61 | 0.33 |
| Jarak Poskesdes | 0.67 | -0.41 | -0.34 | 0.12 | -0.40 | 0.29 | -0.43 | -0.48 | -0.33 | -0.43 | 1.00 | 0.42 | -0.47 | -0.40 |
| Jarak Polindes | 0.10 | -0.04 | 0.15 | 0.26 | -0.02 | 0.17 | -0.02 | -0.03 | 0.41 | -0.04 | 0.42 | 1.00 | -0.02 | -0.17 |
| Jarak Apotek | -0.75 | 0.89 | 0.46 | 0.48 | 0.91 | 0.03 | 0.92 | 0.97 | 0.23 | 0.61 | -0.47 | -0.02 | 1.00 | 0.79 |
| Jarak Toko Obat Jamu | -0.55 | 0.76 | 0.38 | 0.41 | 0.79 | 0.08 | 0.84 | 0.86 | 0.18 | 0.33 | -0.40 | -0.17 | 0.79 | 1.00 |

Gambar 4.2. Korelasi antara kepadatan Penduduk dengan jarak terdekat menuju sarana kesehatan

Matriks korelasi pada Gambar 4.1 menunjukkan adanya hubungan negative antara kepadatan penduduk dengan jumlah sarana Kesehatan yaitu puskesmas rawat inap, puskesmas pembantu, praktek bidan, poskesdes, polindes dan toko obat jamu. Hubungan linear negative antara kepadatan penduduk dengan jumlah sarana Kesehatan dapat menunjukkan bahwa terdapat kepadatan penduduk yang cukup tinggi tetapi memiliki jumlah sarana Kesehatan cenderung sedikit, atau sebaliknya.

Selanjutnya, matriks korelasi pada Gambar 4.2 menunjukkan hanya terdapat dua sarana Kesehatan yang memiliki hubungan positif antara jarak terdekat menuju sarana Kesehatan dengan kepadatan penduduk yaitu poskesdes dan polindes. Selain itu, jarak menuju puskesmas pembantu menunjukkan nilai 0 yang artinya peubah tersebut tidak memiliki hubungan linear dengan kepadatan penduduk. Sebaran puskesmas pembantu di Jawa Barat dapat dilihat pada Gambar 4.3 berikut:

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana



Gambar 4.3. Sebaran Puskesmas Pembantu Kab/kota di Jawa Barat

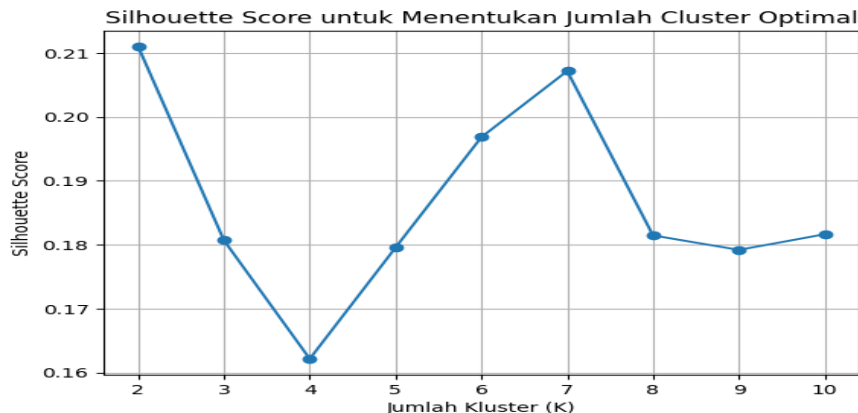
Gambar 4.3 menunjukkan adanya hubungan *negative* antar kepadatan penduduk, jumlah puskesmas pembantu dan jarak menuju puskesmas pembantu. Hal ini terjadi karena terdapat wilayah kota/kabupaten memiliki jumlah penduduk yang tinggi tetapi tidak sebanding dengan jumlah fasilitas puskesmas pembantu, sehingga jarak menuju fasilitas kesehatan tersebut menjadi semakin jauh. Seperti kota Bandung memiliki kepadatan penduduk yang cukup tinggi tetapi menjadi salah satu daerah dengan jumlah puskesmas pembantu lebih sedikit sehingga jarak untuk menuju sarana Kesehatan tersebut semakin jauh. Di sisi lain, Kabupaten Sukabumi memiliki banyak puskesmas pembantu sedangkan kepadatan penduduk di daerah tersebut cenderung rendah.

4.2 Silhouette Score

Silhouette Score digunakan untuk melihat kualitas dan kekuatan cluster, seberapa baik atau buruknya suatu obyek ditempatkan dalam suatu cluster. Metode ini merupakan gabungan dari metode separasi dan kohesi [16]. Penentuan jumlah *cluster* dengan *silhouette score* yang membedakan penelitian ini dengan penelitian [7] yang menentukan jumlah *cluster* dengan *pseudo-F*

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI

Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana

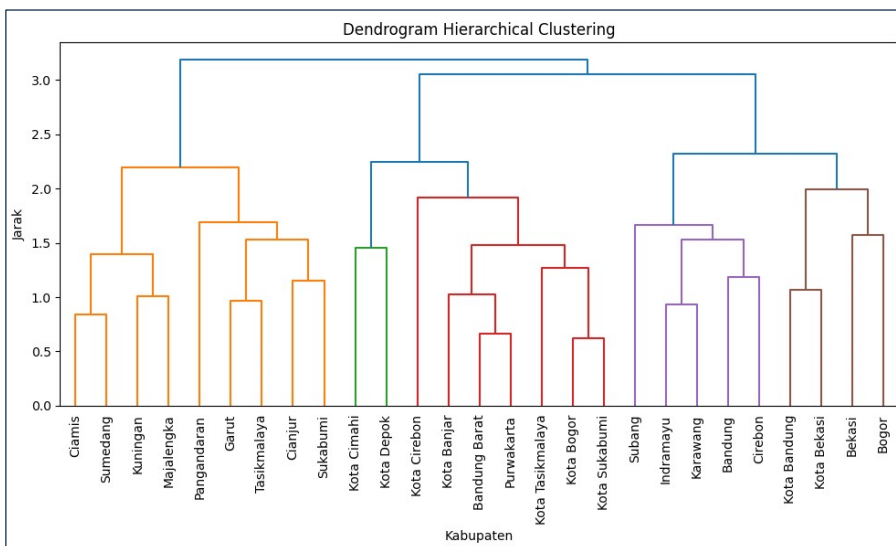


Gambar 4.4. Jumlah *cluster* optimal

Pada pengelompokan data menggunakan *koefisien silhouette* sebagai metode penentu dalam jumlah *cluster* optimal pada *clustering agglomerative* menunjukkan hasil berupa jumlah *cluster* optimal sebanyak 2 *cluster*. Hal tersebut terlihat karena nilai *silhouette* tertinggi terletak pada jumlah *cluster* sebanyak 2 *cluster* dengan nilai rata-rata *silhouette* sebesar 0.21. Nilai *silhouette* tertinggi menunjukkan bahwa objek tersebut terisolasi dengan baik dari *cluster* lain karena objek-objeknya memiliki karakteristik yang serupa, sehingga terdapat dua *cluster* optimal yang akan digunakan untuk mengelompokkan sarana kesehatan di Jawa Barat.

4.3 Complete Linkage Clustering

Penelitian ini memanfaatkan *complete linkage clustering* yaitu salah satu metode dalam *hierarchical clustering* yang menghitung jarak antara dua *cluster* dengan cara mencari jarak maksimum antara setiap data *point* dan *cluster* pertama dengan jumlah data *point* dalam *cluster* kedua. Metode ini termasuk dalam pendekatan *agglomerative*, dimana setiap elemen awalnya berbeda dalam *cluster* sendiri dan kemudian digabungkan berdasarkan jarak antar *cluster*. Gambar 4.5 menunjukkan *dendrogram* yang terbentuk dari *complete linkage clustering*.



Gambar 4.5. Dendrogram *complete linkage clustering*

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana

Gambar 4.5 menunjukkan *cluster* yang menghasilkan dendrogram yang cukup kompleks, dimana setiap setiap *cluster* memiliki partisi yang cukup jelas. Hal ini mempermudah saat memotong dendrogram untuk menentukan banyaknya *cluster*. Gambar tersebut menunjukkan bahwa kabupaten dan kota di Jawa Barat dapat dikelompokkan menjadi lima kelompok besar dan dua kelompok spesifik. Kedekatan antar objek di dendrogram diukur dengan menggunakan jarak. Jarak yang lebih kecil menunjukkan bahwa dua objek lebih mirip satu sama lain.

4.4 Agglomerative Nesting

Metode *Agglomerative Nesting* (AGNES) adalah pendekatan dalam analisis *cluster* yang mengelompokkan data kedalam *cluster-cluster* berdasarkan tingkat kesamaan antar observasi. Pembagian *cluster* sarana Kesehatan di Jawa Barat tersaji pada tabel 2 berikut:

Tabel 4.2. Nama Kab/kota Hasil Cluster

| Cluster 1 | | Cluster 2 |
|-----------------|---------------------|----------------|
| 0 Bandung | 12 Kota Bekasi | 4 Ciamis |
| 1 Bandung Barat | 13 Kota Bogor | 5 Cianjur |
| 2 Bekasi | 14 Kota Cimahi | 7 Garut |
| 3 Bogor | 15 Kota Cirebon | 19 Kuningan |
| 6 Cirebon | 16 Kota Depok | 20 Majalengka |
| 8 Indramayu | 17 Kota Sukabumi | 21 Pangandaran |
| 9 Karawang | 18 Kota Tasikmalaya | 24 Sukabumi |
| 10 Kota Bandung | 22 Purwakarta | 25 Sumedang |
| 11 Kota Banjar | 23 Subang | 26 Tasikmalaya |

Cluster yang terbentuk terdiri dari 18 Kabupaten Kota pada *cluster* 1 dan 9 Kabupaten Kota pada *cluster* 2. Karakteristik kedua *cluster* dapat dilihat pada tabel berikut:

Tabel 4.3. Karakteristik *cluster*

| Peubah | Cluster 1 | | | Cluster 2 | | |
|---------------------------------|-----------|-----|-------|-----------|-----|------|
| | mean | min | max | mean | min | max |
| Kepadatan Penduduk | 5447.89 | 849 | 14630 | 791.56 | 423 | 1095 |
| Jumlah RS | 19.83 | 4 | 61 | 6.56 | 2 | 12 |
| Jumlah RS Bersalin | 3.22 | 0 | 21 | 0.22 | 0 | 1 |
| Jumlah Puskesmas Rawat Inap | 11.39 | 0 | 38 | 21.33 | 6 | 42 |
| Jumlah Puskesmas non Rawat Inap | 32.11 | 8 | 85 | 24 | 5 | 41 |
| Jumlah Puskesmas Pembantu | 37.83 | 2 | 108 | 108 | 48 | 190 |
| Jumlah Poliklinik | 161.17 | 16 | 540 | 69.11 | 27 | 121 |
| Jumlah Praktek Dokter | 218.22 | 32 | 558 | 130.56 | 41 | 185 |
| Jumlah Rumah Bersalin | 23.33 | 0 | 130 | 8.89 | 0 | 21 |
| Jumlah Praktek Bidan | 404.78 | 31 | 1042 | 512.67 | 163 | 808 |
| Jumlah Poskesdes | 40.22 | 0 | 297 | 159.44 | 64 | 289 |
| Jumlah Polindes | 19.06 | 0 | 132 | 26.67 | 0 | 115 |
| Jumlah Apotek | 197.44 | 34 | 394 | 141.33 | 47 | 197 |
| Jumlah Toko Obat Jamu | 175.50 | 18 | 569 | 87.44 | 25 | 141 |

JURNAL MATEMATIKA, STATISTIKA DAN KOMPUTASI
Megawati, Nadira Nisa Alwani, Anwar Fitrianto, Erfiani, Alfa Nugraha Pradana

| | | | | | | |
|--------------------------------|-------|------|-------|-------|-------|-------|
| Jarak RS | 5.79 | 0.60 | 14.16 | 19.68 | 8.75 | 28.70 |
| Jarak RS Bersalin | 21.20 | 2.45 | 99.00 | 58.41 | 30.90 | 93.61 |
| Jarak Puskesmas Rawat Inap | 5.07 | 1.93 | 11.38 | 6.93 | 4.86 | 9.80 |
| Jarak Puskesmas non Rawat Inap | 2.82 | 0.08 | 8.26 | 8.29 | 3.58 | 13.62 |
| Jarak Puskesmas Pembantu | 3.64 | 1.13 | 7.33 | 3.14 | 2.23 | 4.78 |
| Jarak Poliklinik | 1.99 | 0.02 | 4.81 | 7.47 | 4.25 | 13.01 |
| Jarak Praktek Dokter | 1.35 | 0.00 | 3.18 | 4.40 | 2.53 | 6.52 |
| Jarak Rumah Bersalin | 19.91 | 1.98 | 99.00 | 57.54 | 21.06 | 99.30 |
| Jarak Praktek Bidan | 0.18 | 0.00 | 0.48 | 0.46 | 0.06 | 0.86 |
| Jarak Poskesdes | 21.93 | 0.17 | 99.80 | 4.09 | 0.72 | 7.53 |
| Jarak Polindes | 33.90 | 3.11 | 99.80 | 51.40 | 5.09 | 98.72 |
| Jarak Apotek | 1.50 | 0.00 | 3.86 | 4.74 | 2.61 | 6.67 |
| Jarak Toko Obat Jamu | 1.81 | 0.00 | 5.86 | 6.00 | 2.54 | 14.98 |

Tabel diatas menjelaskan bahwa kepadatan penduduk di wilayah kabupaten/kota pada *cluster* satu memiliki jumlah yang lebih tinggi dibanding *cluster* dua. Rata-rata kepadatan penduduk di *cluster* satu sebesar 5447.89 jiwa/km² sedangkan *cluster* dua hanya memiliki rata-rata kepadatan penduduk sebesar 791.56 jiwa/km². Perbedaan kepadatan penduduk ini tercermin dalam jumlah fasilitas kesehatan terutama rumah sakit.

Cluster satu memiliki rata-rata jumlah rumah sakit sebanyak 19.83 unit, sedangkan *cluster* dua hanya memiliki rata-rata rumah sakit sebanyak 6.57 unit. Hal ini berdampak pada jarak yang harus ditempuh penduduk untuk mengakses fasilitas kesehatan. Penduduk pada *cluster* satu hanya perlu menempuh jarak rata-rata 5.79 km untuk mencapai rumah sakit, sedangkan penduduk pada *cluster* dua harus menempuh jarak rata-rata 19.68 km. Di sisi lain, *cluster* dua memiliki jumlah puskesmas rawat inap dan puskesmas pembantu yang lebih tinggi dibandingkan *cluster* satu, hal ini dapat menjadi alternatif dari jarak tempuh yang lebih jauh jika ke rumah sakit.

Berdasarkan hasil tersebut, diketahui bahwa *cluster* satu memiliki akses kesehatan yang lebih baik terhadap layanan kesehatan rujukan, yaitu rumah sakit. Hal ini dikarenakan kepadatan penduduk di *cluster* satu yang lebih tinggi. Sedangkan *cluster* dua memiliki akses yang lebih baik terhadap layanan kesehatan primer, yaitu puskesmas. Hal ini dikarenakan jumlah puskesmas di *cluster* dua lebih tinggi.

5. KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan, pemetaan sarana kesehatan kabupaten/kota di Provinsi Jawa Barat menggunakan analisis *Cluster Agglomerative Nesting* menunjukkan adanya dua *cluster* optimal. Penentuan *cluster* optimal dilakukan dengan metode *silhouette score*, sedangkan jarak antar dua *cluster* diukur menggunakan algoritma *complete linkage clustering*. Namun anggota kedua *cluster* tersebut tidak seimbang. *Cluster* pertama terdiri dari 18 kabupaten/kota dengan rata-rata kepadatan penduduk yaitu 5447.89 jiwa/km², dan *cluster* dua terdiri dari 9 kabupaten/kota dengan rata-rata kepadatan penduduk yaitu 791.56 jiwa/km². Perbedaan kepadatan penduduk yang cukup besar menyebabkan adanya perbedaan kebutuhan akan fasilitas kesehatan pada masing-masing *cluster*

REFERENSI

- [1]. Agresti, A., 2002. *Categorical Data Analysis Second Edition*, John Willey and Sons, New York.
- [2]. Andina, K., 2017. Klasifikasi Informasi, Anjuran dan Larangan pada Hadits Shahih Bukhari menggunakan Metode Support Vector Machine, *e-Proceeding of Engineering*, p. 5014.
- [3]. Anggraeni, M. R., Yudatama, U. & M., 2023. Clustering Prevalensi Stunting Balita Menggunakan Hierarchical Clustering, *Jurnal Media Informatika Budidarma*, pp. 351-359.
- [4]. Azzahra, A. & Wijayanto, A. W., 2022. Perbandingan Agglomerative Hierarchical dan K-Means dalam Pengelompokan Provinsi Berdasarkan Pelayanan Kesehatan Maternal, *Jurnal Sistem Informasi*, Vol. 11 No. 2, pp. 481-495.
- [5]. Dewi, A. F. & Ahadiyah, K., 2022. Agglomerative Hierarchy Clustering Pada Penentuan Kelompok Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Pendidikan, *Zeta - Math Journal*, vol 7 No. 2.
- [6]. Goreti, M., Novia, Y. & Wahyuningsih, S., 2016. Perbandingan Hasil Analisis Cluster dengan Menggunakan Metode Single Linkage dan Metode C-Means, *Ekspensial*, pp. 9-16.
- [7]. Hajar, M., Nasution, Y. N. & Prangga, S., 2022. Penerapan Metode Agglomerative Nesting (AGNES) pada Pengelompokan Wilayah Desa Atau Kelurahan Di Kabupaten Kutai Kartanegara, *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*.
- [8]. Hosmer, L. S. & D. W., 2000. *Applied Logistic Regression*, New York: John Wiley & Sons.
- [9]. Irawan, I., Rini, D. P. & E., 2021. Peningkatan Performa Algoritma CART Dengan Seleksi Fitur Menggunakan ABC Untuk Penilaian Kredit, *Teknik Informatika dan Sistem Informasi*, vol. 8, pp. 199-208.
- [10]. Kasoqi, I. A., Hayati, M. N. & Goejantoro, R., 2021. Pengelompokan Desa Atau Kelurahan Di Kutai Kartanegara Menggunakan Algoritma Divisive Analysis, *J. Stat. Univ. Muhammadiyah Semarang*, vol.9, no. 2, p. 101, doi: 10.26714/jsunimus.9.2.2021.101-108.
- [11]. Kemenkes, RI., 2022. Sehat Negeriku, Kementerian Kesehatan, 29 September [Online]. Available: <https://sehatnegeriku.kemkes.go.id/baca/rilismedia/20220929/0541166/penyakit-jantung-penyebab-utama-kematian-kemenkes-perkuat-layanan-primer/>.
- [12]. Li, T., Rezaipnah, A. & Din, E. M. T. E., 2022. An ensemble agglomerative hierarchical clustering algorithm based on clusters clustering technique and the novel similarity measurement, *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no.6, pp. 3828-3842
- [13]. Maghfiroh W., Ulinuha, N. & Fanani, A., 2019. Penerapan Fuzzy C-Means dalam Mengelompokkan Kabupaten/Kota Berdasarkan Fasilitas Pelayanan Kesehatan di Jawa Timur, *Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol.4 No.1.
- [14]. Nahar, J., 2017. Penerapan Metode Multidimensional Scaling dalam Pemetaan Sarana Kesehatan di Jawa Barat, *Matematika Integratif*, pp. 43-50.
- [15]. Pratikto, R. O. & Damastuti, N., 2021. Klasterisasi Menggunakan Agglomerative Hieraarchical Clustering untuk Memodelkan Wilayah Banjir, *Journal of Information Technology and Computer Science*, Vol. 6 No.1, pp. 13-20.
- [16]. Putu, N., Merliana, E., & Santoso, A. J., 2018. Analisa penentuan jumlah cluster terbaik pada metode KMeans, *Seminar Nasional Multi Disiplin Ilmu*, pp. 978-979.
- [17]. Rian Dani, A. T., Wahyuningsih, S. & N. A. Rizki., 2019. Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu, *Jambura Journal of Mathematics*.

- [18]. Rizky, A. M. & Y., 2019. Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu., *Jambura Journal of Matheatics*, pp. 64-78.
- [19]. Ros, F. & Guillaume, S., 2019. A hierarchical clustering algorithm and an improvement of the single linkage criterion to deal with noise, *Expert Systems with Applications*, vol. 128, pp. 96-108.
- [20]. Santosa, B., 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- [21]. Tan, P., Steinbach, M. & Kumar, V., 2006. *Introduction to Data Mining*. USA: Person Education, Inc.
- [22]. Tokuda, E. K., Comin, C. H. & Costa, L. d. F., 2022. Revisiting Agglomerative clustering, *Physica A: Statistical Mechanics and its Applications*.
- [23]. Widodo, E., Ermayani, P., Laila, L. N. & Madani, A. T., 2021. Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Analisis Hierarchical Agglomerative Clustering, *Seminar Nasional Official Statistics*.