

Analisis Cluster Hierarki untuk Pengelompokan Provinsi di Indonesia berdasarkan Jumlah *Base Transceiver Station* dan Kekuatan Sinyal

T Apriliana¹, E Widodo²

^{1,2}Program Studi Statistika, Universitas Islam Indonesia

E-mail: 20611088@students.uui.ac.id¹, 966110103@uui.ac.id²

Abstrak. Pembangunan akses internet dan infrastruktur telekomunikasi sangat penting untuk memfasilitasi interkoneksi antara perangkat atau sistem. Pembangunan akses internet dan infrastruktur telekomunikasi yang mencukupi menyebabkan interkoneksi antara perangkat dapat terwujud dengan baik dan memudahkan pertukaran informasi dan data yang lebih efektif. Adanya *Base Transceiver Station* (BTS) dan penerimaan sinyal telepon seluler di suatu wilayah sangat berpengaruh pada kemudahan komunikasi dan bertukar informasi, di mana ini merupakan aspek yang dapat mengukur perkembangan teknologi di wilayah tersebut. Tujuan dalam penelitian ini yaitu mengelompokkan provinsi di Indonesia dengan parameter yang digunakan adalah jumlah keberadaan BTS, kekuatan sinyal telepon baik lemah dan kuat, serta kekuatan sinyal internet 4G dengan menggunakan analisis *cluster* hierarki. Hasil penelitian menunjukkan bahwa terdapat 3 kelompok berdasarkan tiga indikator yang digunakan, dengan kelompok yang memiliki akses internet paling rendah terdiri dari 19 provinsi, sedangkan kelompok yang memiliki akses internet paling tinggi terdiri dari 3 provinsi yang berada di pulau Jawa. Hasil dari penelitian ini diharapkan dapat digunakan sebagai dasar acuan pertimbangan bagi program pemerintah untuk perencanaan dan evaluasi terkait kelompok wilayah yang masih memerlukan pembangunan dan perbaikan infrastruktur telekomunikasi di Indonesia.

Kata kunci: analisis klaster hierarki; *Base Transceiver Station* (BTS); sinyal telepon; internet 4G.

Abstract. The development of internet Access and telecommunication infrastructure is crucial to facilitating interconnectivity between devices or systems. Sufficient internet access and telecommunication infrastructure enable good interconnectivity between devices and facilitate a more effective exchange of information and data. The accessibility and quality of communication and information exchange, which are key indicators of technological advancement, can be significantly influenced by the existence of a *Base Transceiver Station* (BTS) and mobile phone signal reception in a particular area. The aim of this research is to cluster Indonesian provinces with the parameters used are the number of BTS presence, weak and strong phone signal strength, and 4G internet signal strength using hierarchical clustering analysis. As a result, showed that there are 3 groups based on three indicators used, the group with the lowest internet access consisting of 19 provinces, while the group with the highest internet access consisted of 3 provinces on the island of Java. The study's findings are anticipated to serve as a basis for planning and assessing

government program targets related to regional groups that still require the development and improvement of telecommunications infrastructure in Indonesia.

Keywords: *hierarchical clustering analysis; Base Transceiver Station (BTS); telephone signal; 4G internet*

1. Pendahuluan

Di tengah upaya pemulihan pandemi saat ini, perkembangan teknologi semakin cepat dan hampir seluruh aspek kehidupan berkaitan dengan teknologi dan informasi. Perkembangan teknologi ini membentuk suatu era baru yang disebut dengan era digital yang menjadi suatu keharusan untuk dijalani pada saat pandemi Covid-19 mewabah seluruh negara di dunia [1]. Pandemi Covid-19 membuat perubahan yang menuntut semua masyarakat baik dalam kehidupan ekonomi, pendidikan maupun sektor lainnya tidak gagap dengan teknologi, terutama teknologi informasi/digital.

Kehadiran dan pemanfaatan teknologi bukan hanya untuk pembangunan tetapi juga untuk mendukung interaksi sosial oleh masyarakat sehari-hari. Kementerian Komunikasi dan Informatika telah menjalankan program transformasi digital dalam beberapa tahun terakhir yang menekankan pada pengembangan infrastruktur untuk memperkecil kesenjangan akses terhadap teknologi digital [2]. Interkonektivitas tidak lepas dari koneksi atau akses internet ke jaringan/perangkat lain karena berkaitan dengan kemampuan perangkat untuk berkomunikasi dan berinteraksi secara efektif antar manusia. Dengan demikian, sangat penting bagi pemerintah untuk memperhatikan infrastruktur akses internet yang memfasilitasi interkonektivitas antara perangkat. Kemajuan perkembangan internet di Indonesia dapat dikaitkan dengan upaya pemerataan infrastruktur jaringan yang dilakukan oleh setiap penyelenggara layanan komunikasi di daerah perkotaan maupun di perdesaan [3]. Dengan teknologi internet, komunikasi antar individu di seluruh dunia dapat terjadi dengan mudah tanpa adanya batasan. Dalam perkembangannya, internet telah bermanfaat bagi banyak berbagai bidang seperti ekonomi, pendidikan, dan komunitas lain baik dalam skala kecil maupun besar.

Pada publikasi Badan Pusat Statistik (BPS) memperoleh data yang menunjukkan adanya peningkatan penggunaan telepon seluler di Indonesia dalam beberapa tahun terakhir [4]. Peningkatan penggunaan telepon seluler dipengaruhi oleh kemajuan sistem komunikasi seluler, termasuk teknologi GSM (*Global System for Mobile Communication*). Sementara pada teknologi GSM, BTS merupakan komponen penting dari sebuah sel *site* yang menyediakan layanan telekomunikasi [5]. Untuk kekuatan sinyal telepon seluler, baik pada keadaan sinyal telepon maupun internet bervariasi dari berbagai provinsi di Indonesia. Dengan demikian, variabel-variabel infrastruktur telekomunikasi seperti keberadaan BTS dan kekuatan sinyal telepon seluler dapat digunakan sebagai parameter untuk pengelompokan.

Analisis *clustering* adalah salah satu metode statistik multivariat terapan untuk mengelompokkan setiap objek yang paling mirip dengan objek lain dalam *cluster* yang sama. Analisis ini digunakan untuk mengelompokkan provinsi di atas [6]. Dalam analisis ini terdapat dua metode, yaitu analisis *cluster* hierarki dan non-hierarki. Metode *cluster* hierarki terdiri dari metode *agglomerative* dan *devisive*. Penggunaan metode *agglomerative* di mana seluruh objek dianggap *cluster* yang digabungkan berdasarkan jarak terdekat antara *cluster*, sehingga membentuk *q cluster* [7]. Tujuan menggunakan analisis *cluster* hierarki pada penelitian ini adalah untuk mengelompokkan provinsi yang memiliki kemiripan karakteristik yang paling dekat ditinjau dari tiga ukuran indikator sebagai parameter penelitian. Indikator-indikator tersebut yaitu keberadaan BTS, kekuatan sinyal telepon baik lemah dan kuat, serta kekuatan sinyal internet 4G yang membentuk hierarki atau tingkatan antar provinsi.

Secara umum, hasil pengelompokan pada penelitian ini dapat menggambarkan kondisi akses internet di seluruh provinsi, di mana hal ini dapat menjadi dasar acuan perencanaan dan evaluasi program pemerintah terkait pemerataan jaringan internet di desa/kelurahan seluruh provinsi di Indonesia. Misalnya dengan melakukan pembangunan BTS di beberapa wilayah di provinsi untuk menghadirkan sinyal 4G atau dilakukan pemeliharaan infrastruktur secara berkala untuk mempertahankan kualitas sinyal telepon dan

jaringan internet. Metode analisis *cluster* hierarki pada penelitian ini menggunakan *single linkage*, *complete linkage*, *average linkage*, dan *ward*. Keuntungan menggunakan analisis *cluster* hierarki adalah dapat disajikan dalam bentuk dendrogram sehingga mudah dipahami dan dapat membantu dalam memilih *cluster* yang optimal untuk data yang dianalisis.

Penelitian terdahulu terkait analisis *cluster* menggunakan metode hierarki antara lain penelitian yang dilakukan oleh [8] mengenai pengelompokan provinsi di Indonesia berdasarkan tingkat kemiskinan menggunakan analisis *hierarchical agglomerative clustering* menghasilkan kesimpulan dengan menggunakan metode terbaik setelah membandingkan koefisien *agglomerative* tingkat kemiskinan tertinggi berada di provinsi Papua [8]. Pada penelitian ini menggunakan metode *single linkage*, *complete linkage*, *average linkage*, dan *ward*. Penelitian lain dilakukan oleh Nurafiza Thamrin dan Arie Wahyu Wijayanto (2021) mengenai analisis *cluster* dengan menggunakan *hard clustering* dan *soft clustering* untuk pengelompokan tingkat kesejahteraan kabupaten/kota di pulau Jawa [9]. Hasil penelitian ini memperoleh bahwa pengelompokan yang dilakukan dengan metode terbaik adalah menggunakan metode *agglomerative ward linkage*. Analisis *cluster* dalam penelitian ini menggunakan *hard clustering* yaitu *k-means*, *k-medoids* dan *hierarchical agglomerative* serta *soft clustering* yaitu *fuzzy c-means*.

2. Metode

2.1. Asumsi Analisis Cluster

Dalam analisis *cluster*, sampel yang representatif dan tidak adanya tanda multikolinieritas antarvariabel merupakan asumsi yang diperlukan [10].

1) Sampel Representatif

Penggunaan uji Kaiser-Mayer-Olkin (KMO) dapat memberikan evaluasi apakah sampel representatif atau tidak. Hal ini penting karena menggunakan sampel yang mewakili populasi dapat memberikan hasil terbaik. Uji KMO merupakan indeks yang digunakan untuk mengukur ketepatan sampel yang sudah mencukupi secara keseluruhan tiap indikator. Asumsi bahwa sampel mewakili populasi terpenuhi atau sampel representatif jika koefisien KMO antara 0,5 dan 1. Persamaan uji KMO adalah sebagai berikut [11].

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad (1)$$

dengan p adalah banyaknya variabel, r_{ij} adalah koefisien korelasi antara variabel i dan j , serta r_{ij} untuk koefisien korelasi antara variabel i dan j .

2) Multikolinieritas

Untuk mengetahui adanya multikolinieritas yaitu dengan melihat koefisien korelasi antar variabel independen menggunakan *Variance Inflation Factor* (VIF). Jika koefisien VIF melebihi 10 maka menunjukkan adanya gejala multikolinieritas sehingga perlu melakukan analisis komponen utama untuk menanganinya.

2.2. Principal Component Analysis (PCA)

PCA adalah teknik analisis dengan mengubah sekelompok variabel yang saling berkorelasi menjadi satu variabel baru yang independen. Tujuan utama analisis PCA adalah untuk mereduksi jumlah variabel tanpa mengurangi karakteristik dari data asli secara signifikan. PCA dihitung berdasarkan koefisien *eigen* dan *eigencektor* untuk menghitung penyebaran data dalam suatu dataset [7]. Penentuan jumlah komponen utama yang terpilih berdasarkan koefisien *eigen* > 1.

Tahapan dalam melakukan analisis PCA dimulai dengan standarisasi data, di mana pada tahap ini bertujuan untuk menghasilkan data interval yang lebih seimbang dengan menormalkan rentang koefisien

pada data. Tahap selanjutnya menentukan hubungan antarvariabel penelitian dengan menghitung matriks varians kovarians. Selanjutnya, peneliti menghitung koefisien *eigen* dan *eigenvector* yang diperoleh dari hasil matriks kovarians. Persamaan berikut untuk menghitung koefisien *eigen* [12].

$$[\lambda I - A] = 0 \quad (1)$$

dengan λ adalah koefisien *eigen*, I untuk matriks identitas, dan A untuk matriks data.

Kemudian dilanjutkan dengan menghitung *eigenvector* menggunakan persamaan berikut dengan X adalah *eigen vector* [12].

$$[\lambda I - A][X] = 0 \quad (2)$$

2.3. Analisis Cluster

Clustering adalah metode yang paling sering digunakan untuk analisis data eksplorasi dan klasifikasi *unsupervised* [13]. Analisis *cluster* merupakan statistika multivariat yang memungkinkan dilakukan pengelompokan berdasarkan kesamaan karakteristiknya. Dalam metode ini, objek yang memiliki karakteristik yang sama dalam satu kelompok memiliki tingkat kesamaan yang tinggi, sedangkan perbedaan karakteristik antar kelompoknya rendah. Tujuan dari analisis *cluster* adalah untuk mengelompokkan objek menjadi beberapa kelompok yang memiliki sifat berbeda antar kelompok, sehingga objek yang berada dalam satu kelompok memiliki karakteristik yang relatif homogen [14]. Analisis *cluster* memiliki dua metode utama, yaitu metode hierarki dengan penentuan banyak kelompok yang terbentuk dilihat berdasarkan hasil dendrogram dan metode non-hierarki dengan penentuan banyak kelompok ditentukan di awal penelitian.

2.4. Metode Hierarki

Analisis *cluster* dengan metode hierarki digunakan sebagai pengelompokan objek secara hierarki berdasarkan kemiripan sifatnya yang belum diketahui jumlah *cluster* yang terbentuk. Metode hierarki digunakan pada data yang memiliki struktur hierarkis atau terdiri dari beberapa sub-kelompok yang saling berkaitan dan umumnya jumlah sampel yang digunakan relatif kecil. Hasil pengelompokan dengan metode hierarki ditampilkan dalam sebuah diagram pohon yang disebut dengan dendrogram.

Terdapat dua pendekatan dasar dalam metode hierarki yaitu agglomeratif (pemusatan) dan *disive* (penyebaran). Pendekatan agglomeratif merupakan tipe yang sering digunakan dalam analisis *cluster* hierarki. Pendekatan *agglomerative* biasanya melibatkan pembentukan *cluster* baru dengan menggabungkan objek individu berdasarkan kedekatannya satu sama lain. Dengan menentukan jarak antar *cluster* maka kedekatan dapat ditentukan [15]. Terdapat beberapa metode *cluster* hierarki dengan pendekatan *agglomeratif* sebagai berikut.

1) *Single Linkage* adalah teknik pengelompokan yang didasarkan pada jarak terdekat antara dua objek. Pada awalnya, pengelompokan *single linkage* memilih jarak terkecil dan menggabungkan objek yang sesuai. Berikut rumus perhitungan jarak yang digunakan.

2)

$$d_{(UV)W} = \min (d_{UW}, d_{VW}) \quad (3)$$

dengan:

$d_{(UV)W}$ = jarak minimum antara kelompok (UV) dan kelompok W

d_{uv} = jarak antara tetangga terdekat dari *cluster* U dan W

d_{vw} = jarak antara tetangga terdekat dari *cluster* V dan W

- 3) *Complate Linkage* adalah cara pengelompokan objek yang memiliki jarak paling jauh atau kesamaan yang sedikit. Rumus perhitungan jarak adalah sebagai berikut.

$$d_{(UV)W} = \max (d_{UW}, d_{VW}) \quad (5)$$

- 4) *Average Linkage* adalah pengelompokan yang dibentuk berdasarkan koefisien rata-rata pada jarak seluruh individu dalam satu kelompok dengan rata-rata jarak seluruh individu pada kelompok lainnya. Berikut ini rumus untuk perhitungan jarak.

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{(UV)}N_W} \quad (4)$$

- 5) Metode *Ward* adalah pengelompokan dengan jarak antar dua kelompok adalah jumlah kuadrat antara dua kelompok untuk seluruh variabel. Metode ini dapat meminimumkan varians dalam kelompok dan cenderung digunakan untuk melakukan kombinasi kelompok-kelompok dengan jumlah yang kecil. Untuk pengelompokan dengan metode *ward*, ukuran yang digunakan adalah *Sum of Square* (SSE).

$$SSE = \sum_{i=1}^N (x_j - \bar{x}) (x_j - \bar{x}) \quad (5)$$

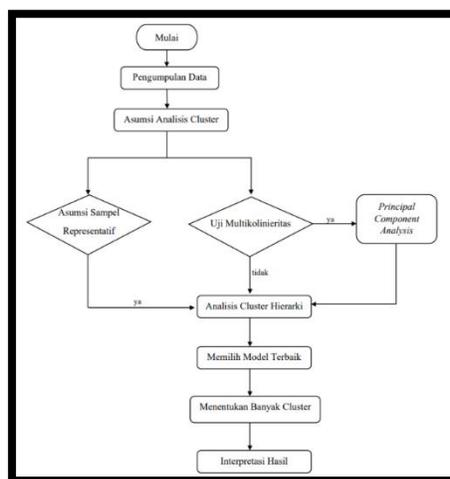
2.5. Data dan Variabel Penelitian

Data diperoleh dari *website* resmi Publikasi Badan Pusat Statistik (BPS), di mana merupakan data sekunder mengenai data akses dan infrastruktur telekomunikasi di Indonesia tahun 2021 [4]. Penelitian ini menggunakan variabel sebagai indikator akses dan infrastruktur telekomunikasi berikut.

- X1: Banyak desa/kelurahan yang memiliki *Base Transceiver Station* (BTS) setiap provinsinya
- X2: Banyak desa/kelurahan yang memiliki sinyal telepon seluler baik sinyal kuat dan lemah
- X3: Banyak desa/kelurahan berdasarkan penerimaan sinyal internet 4G

2.6. Tahapan Analisis Data

Berikut adalah diagram alir yang dilakukan dalam penelitian ini di mana menunjukkan tahapan-tahapan dalam analisis.



Gambar 1. Tahapan Analisis Penelitian

- 1) Dimulai dengan pengumpulan data
Pada penelitian menggunakan data akses dan infrastruktur telekomunikasi di Indonesia tahun 2021 yang diperoleh dari *website* resmi pada Publikasi Badan Pusat Statistik (BPS).
- 2) Asumsi analisis *cluster*
Dalam analisis *cluster* asumsi yang harus terpenuhi antara lain adalah sampel harus representatif dan bebas dari multikolinieritas. Apabila pada data terjadi multikolinieritas, maka dapat dilanjutkan dengan analisis komponen utama yang mereduksi variabel tanpa menghilangkan informasi yang memuat data aslinya.
- 3) Analisis *cluster*
Dilakukan pengelompokan provinsi di Indonesia berdasarkan indikator keberadaan BTS, kekuatan sinyal telepon, dan kekuatan sinyal internet 4G menggunakan analisis *cluster* hierarki. Metode pada analisis *cluster* hierarki yang digunakan antara lain metode *single linkage*, *complete linkage*, *average linkage*, dan *ward*.
- 4) Pilih model terbaik
Peneliti membandingkan hasil pengelompokan pada 4 metode yang telah dilakukan dengan menggunakan koefisien *agglomerative* yang paling tinggi sebagai model terbaik. Metode dengan model terbaik akan ditetapkan untuk dibentuk sebagai dendogram.
- 5) Menentukan banyaknya *cluster*
Penentuan banyak *cluster* dilakukan dengan dendogram didasarkan pada jarak yang paling besar [16].

3. Hasil dan Pembahasan

3.1. Uji Asumsi Cluster

3.1.1 Uji Kaiser-Meyer-Olkin (KMO)

Uji KMO sebagai uji yang mengukur kecukupan pengambilan sampel secara keseluruhan dan mengukurnya untuk setiap indikator. Dalam hal ini, apabila koefisien KMO yang diperoleh lebih dari 0,5 maka asumsi sampel yang mewakili populasi terpenuhi atau sampel representatif. Hasil koefisien KMO dapat dilihat pada Tabel 1 berikut ini.

Tabel 1. Koefisien *Kaiser-Mayer-Olkin*

Kaiser Mayer Olkin (KMO)	0,66
-----------------------------	------

Berdasarkan Tabel 1. di atas, koefisien KMO yang diperoleh pada yaitu sebesar 0,66 yang menunjukkan bahwa koefisien KMO yang diperoleh melebihi 0,5 sehingga sampel mewakili populasi atau sampel representatif. Dengan demikian, sampel dapat digunakan untuk analisis lebih lanjut.

3.1.2 Uji Multikolinieritas

Uji multikolinieritas merupakan suatu kondisi yang menunjukkan adanya hubungan kuat antara dua variabel bebas. Pengujian ini dilakukan dengan membandingkan koefisien *Variance Inflation Factor* (VIF). Jika koefisien VIF melebihi 10 maka menunjukkan adanya multikolinieritas. Tabel 2 menunjukkan hasil koefisien VIF sebagai berikut.

Tabel 2. Koefisien *Variance Inflation Factor*

X1	X2	X3
11,83942	40,02016	64,90653

Berdasarkan Tabel 2 di atas, koefisien VIF lebih dari 10 sehingga menunjukkan bahwa semua variabel mengandung multikolinieritas atau terjadi korelasi antarvariabel.

3.2. Principal Component Analysis (PCA)

Principal component analysis (PCA) adalah langkah yang tepat digunakan dalam analisis *cluster* apabila terdapat multikolinieritas. Tujuan melakukan PCA untuk mereduksi variabel yang ada menjadi lebih sedikit tanpa harus menghilangkan informasi yang memuat data aslinya. Untuk menentukan banyak komponen utama yang dihasilkan, dapat dilihat dengan koefisien *eigen* yang menunjukkan lebih dari 1 sebagai berikut.

Tabel 3. Koefisien *Eigen*

X1	X2	X3
2,9035	0.0870	0,0095

Tabel 3 di atas menunjukkan koefisien *eigen* untuk setiap faktornya. Berdasarkan hasil yang diperoleh, pada komponen 1 memiliki koefisien *eigen* lebih dari 1. Hal ini menunjukkan analisis PCA membentuk sebanyak 1 komponen utama di mana koefisien *eigen* yang diperoleh sebesar 2,9035. Tabel 4 di bawah ini merupakan tabel hasil dari analisis PCA.

Tabel 4. Koefisien Hasil PCA

No.	PCA	No.	PCA
1	0,36676082	18	-0,19149287
2	1,21900611	19	-0,05110363
3	-0,17249903	20	-0,15680847
4	0,09424053	21	-0,49126519
5	-0,33105629	22	-0,23691292
6	0,29078547	23	-0,42189639
7	-0,57797620	24	-0,77039204
8	0,20820356	25	-0,36821815
9	-0,66138392	26	-0,36574069
10	-0,69524250	27	0,44025872
11	-0,74726911	28	-0,43263204
12	2,85082459	29	-0,69937160
13	2,88468317	30	-0,75470148
14	-0,66138392	31	-0,46731644
15	3,03663388	32	-0,58953766
16	-0,03541306	33	0,59201512
17	-0,47392299	34	-0,44584514

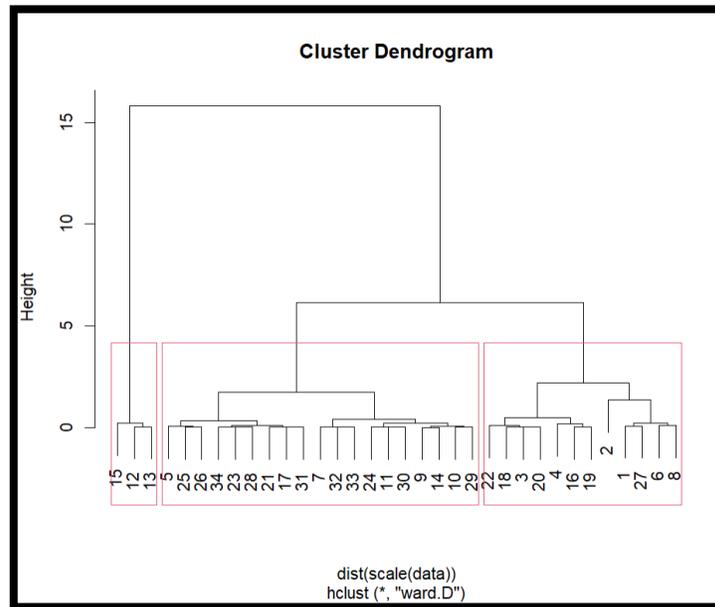
3.3. Analisis Cluster dengan Metode Hierarki

Analisis *cluster* dengan metode hierarki dikenal sebagai dendogram atau pohon, di mana analisis ini dilakukan secara bertahap dan bertingkat. Dalam menentukan banyak *cluster* pada metode ini, diperoleh dengan urutan pasti yang menggabungkan atau membagi *cluster*. Terdapat 4 metode *agglomerative* antara lain metode *single linkage*, *complete linkage*, *average linkage*, dan *ward*. Untuk penentuan metode terbaik dari 4 metode tersebut, analisis *cluster* hierarki ditentukan dengan koefisien agglomeratif. Hasil penentuan dari koefisien agglomeratif dilihat berdasarkan Tabel 5 berikut.

Tabel 5. Koefisien Agglomeratif

Metode	Koefisien Agglomeratif
<i>Single linkage</i>	0,9692635
<i>Complete linkage</i>	0,9776254
<i>Average linkage</i>	0,9764644
<i>Ward</i>	0,9894966

Tabel 5 di atas menunjukkan bahwa metode hierarki yang memperoleh *cluster* terbaik adalah dengan metode *ward* yang ditunjukkan dengan perbandingan koefisien agglomeratif lebih besar dengan metode hierarki lainnya. Dendrogram dengan metode *ward* pada Gambar 2 berikut.



Gambar 2. Dendrogram Metode Ward

Berdasarkan dendrogram yang diperoleh dari metode *ward*, dapat dilihat bahwa objek terbagi dengan sangat baik. di mana untuk setiap *cluster*, terlihat koefisien *height* yang rendah dikarenakan cara kerja metode *ward* dengan meminimumkan koefisien *within sum of squared* (WSS) tiap *cluster*. Pada dendrogram di atas pada *cluster* 1 terdiri dari 12 provinsi, *cluster* 2 terdiri dari 19 provinsi, dan pada *cluster* 3 terdiri dari 3 provinsi.

Untuk mengetahui anggota kelompok yang terbentuk dalam setiap *cluster* menggunakan metode *ward*, ditampilkan dengan Tabel 6 berikut.

Tabel 6. Hasil Pengelompokan dengan Metode Ward

Cluster	Jumlah Anggota	Provinsi
1	12	Aceh, Sumatera Utara, Sumatera Barat, Sumatera Selatan, Lampung, Riau, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Selatan, dan Sulawesi Selatan.
2	19	Bengkulu, Kep. Bangka Belitung, Kepulauan Riau, Jambi, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua.
3	3	Jawa Barat, Jawa Tengah, dan Jawa Timur.

Dari analisis *cluster* dengan menggunakan algoritma hierarki pada metode *ward*, hasil menunjukkan bahwa pada *cluster* 1 terdiri dari 12 provinsi yang memiliki akses dan infrastruktur

telekomunikasi di Indonesia tahun 2021 yaitu provinsi Aceh, Sumatera Utara, Sumatera Barat, Riau, Sumatera Selatan, Lampung, Banten, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Selatan, dan Sulawesi Selatan. Pada *cluster* 2 terdapat 19 provinsi yaitu Jambi, Bengkulu, Kep. Bangka Belitung, Kepulauan Riau, DKI Jakarta, DI Yogyakarta, Bali, Kalimantan Tengah, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat, dan Papua. Sedangkan pada *cluster* 3 terdapat 3 provinsi yaitu Jawa Barat, Jawa Tengah, dan Jawa Timur.

Setelah mendapatkan hasil pengelompokan dengan algoritma hierarki, maka dilanjutkan dengan melakukan profilisasi data pada kelompok yang terbentuk. Hasil profilisasi data dengan menggunakan metode *ward* dapat di lihat pada Tabel 7 berikut ini.

Tabel 7. Profilisasi *Cluster* Terbaik

<i>Cluster</i>	X1	X2	X3
1	1328	2869,583	2167,917
2	476,6842	1131,263	791,1579
3	4689,667	7669,667	6959,667

Berdasarkan Tabel 7 di atas diperoleh profilisasi yang memiliki arti bahwa untuk setiap *cluster* memiliki karakteristik yang berbeda. Pada *cluster* 1 dengan *cluster* lainnya memiliki perbedaan koefisien rata-rata *cluster* yang ditunjukkan dengan perbedaan warna yang mencolok. Warna hijau menunjukkan bahwa keberadaan BTS, penerimaan sinyal telepon seluler yang baik dan sinyal internet 4G di provinsi *cluster* tersebut rendah, warna kuning menunjukkan bahwa keberadaan BTS, penerimaan sinyal telepon seluler yang baik dan sinyal internet 4G di provinsi *cluster* tersebut sedang, dan warna merah menunjukkan bahwa keberadaan BTS, penerimaan sinyal telepon seluler yang baik dan sinyal internet 4G di provinsi *cluster* tersebut tinggi.

Cluster 1 merupakan kelompok provinsi yang memiliki BTS dan penerimaan sinyal telepon serta internet 4G termasuk dalam kategori sedang. Anggota provinsi pada *cluster* 1 terdapat 12 provinsi di mana provinsi terbanyak ada di pulau Sumatera. *Cluster* 2 merupakan kelompok provinsi yang memiliki BTS dan penerimaan sinyal telepon serta internet 4G termasuk dalam kategori rendah. Anggota provinsi pada *cluster* 2 ini jumlah keanggotaannya lebih banyak dibandingkan dengan *cluster* lainnya, yaitu sebanyak 19 provinsi. Sementara itu, *cluster* 3 merupakan kelompok provinsi yang memiliki BTS dan penerimaan sinyal telepon serta internet 4G termasuk dalam kategori tinggi. Anggota pada *cluster* 3 terdiri dari provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur.

4. Kesimpulan

Dengan memperhatikan kuantitas *Base Transceiver Station* (BTS), kualitas sinyal telepon seluler, dan ketersediaan internet 4G, dengan menggunakan analisis klaster dalam penelitian ini bertujuan mengelompokkan provinsi di Indonesia. Metode yang digunakan dalam analisis *clustering* adalah *cluster* hierarki dengan menggunakan variabel jumlah adanya BTS, jumlah desa dengan sinyal telepon baik yang kuat maupun yang lemah, dan jumlah desa dengan sinyal internet 4G.

Dalam analisis ini, dibentuk tiga 3 *cluster* yaitu wilayah dengan BTS, sinyal telepon seluler dan internet 4G baik/tinggi (*C3*) terdiri dari 3 provinsi, wilayah dengan BTS, sinyal telepon seluler dan internet 4G sedang (*C1*) terdiri dari 19 provinsi, dan wilayah dengan BTS, sinyal telepon seluler dan internet 4G rendah (*C2*) terdiri dari 19 provinsi. Berdasarkan hasil *clustering* diperoleh bahwa pada (*C2*) yang terdiri dari 19 provinsi yang kelompok provinsi ini perlu lebih diperhatikan oleh pemerintah untuk pembangunan akses dan infrastruktur telekomunikasi. Hasil pengelompokan ini selanjutnya dapat dijadikan rekomendasi bagi pemerintah tidak hanya untuk wilayah yang memerlukan pemerataan pembangunan akses dan infrastruktur telekomunikasi di Indonesia, tetapi juga dapat dilakukan dengan pemeliharaan infrastruktur dalam proses pemerataan kualitas jaringan internet [17].

Referensi

- [1] K. Hadiono, H. Murti and R. C. N. Santi, "Transformasi Digital di Masa Pandemi Covid-19," *Proceeding SENDIU*, 2021.
- [2] Kominfo, "DEWG Presidensi G20 Indonesia: Konektivitas Digital dan Pemulihan Pascapandemi Covid-19 Perlu Prinsip Inklusif, Memberdayakan, dan Berkelanjutan," Kominfo, 2 2022. [Online]. Available: https://www.kominfo.go.id/content/detail/39939/siaran-pers-no-49hmkominfo022022-tentang-dewg-presidensi-g20-indonesia-konektivitas-digital-dan-pemulihan-pascapandemi-covid-19-perlu-prinsip-inklusif-memberdayakan-dan-berkelanjutan/0/siaran_pers. [Accessed 2 4 2023].
- [3] T. P. P. SDPPI, "Analisis Industri Telekomunikasi Indonesia untuk Mendukung Efisiensi," *Puslitbang Sumber Daya, Perangkat, dan Penyelenggaraan Pos dan Informasi*, pp. 1-61, 2018.
- [4] BPS, "Statistik Telekomunikasi Indonesia 2021," Badan Pusat Statistik, Jakarta, 2022.
- [5] S. Budiyanto and A. Saputra, "Optimalisasi Kinerja (Internet Protocol) IP Clock pada Jaringan Base Transceiver Station (BTS)," *Junal Teknologi Elektro Universitas Mercu Buana*, vol. 7, no. 2, pp. 114-122, 2016.
- [6] S. I. Pratiwi, T. Widiariyah and A. R. Hakim, "Analisis Kluster Metode Ward dan Average Linkage dengan Validasi Dunn Index dan Koefisien Korelasi Cophenitic," *Jurnal Gaussin*, vol. 8, no. 4, pp. 486-495, 2019.
- [7] R. A. Johnson, *Applied Multivariate Statistical Analysis Sixth Edition*, New Jersey: Prentice Hal, Inc, 2007.
- [8] E. Widodo, P. Ermayani, L. N. Laila and A. T. Madani, "Pengelompokan Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Analisis Hierarchical Agglomerative Clustering," *Seminar Nasional Official Statistics 2021*, vol. 2021, no. 1, 2021.
- [9] N. Thamrin and A. W. Wijayanto, "Comparison of Soft and Hard Clustering: A Case Study on Welfare Level in Cities on Java Island," *Indonesian Journal of Statistics and Its Applications*, vol. 5, no. 1, pp. 141-160, 2021.
- [10] J. F. Hair, W. C. Black, B. J. Babin and R. E. Anderson, *Multivariate Data Analysis*, 7 ed., England: Pearson Education Limited.
- [11] A. Muhtadan, "Clustering of District or City in Central Java Based Covid-19 Case using K-Means Clustering," *Jurnal Matematika, Statistika, & Komputasi*, 2020.
- [12] Muhtadi, "Penerapan Principal Component Analysis (PCA) dalam Algoritma K-Means untuk Menentukan Centroid pada Clustering," *Jurnal of Mathematic Teaching*, 2017.
- [13] D. Bakkelund, "Order Preserving Hierarchical Agglomerative Clustering," *Machine Learning*, vol. 111, no. 5, pp. 1851-1901, 2022.
- [14] M. W. Talakua, Z. A. Leleury and A. W. Taluta, "Pengelompokan Kabupaten/Kota di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014," *BAREKENG: Jurnal Matematika dan Ilmu Terapan*, vol. 11, no. 2, pp. 119-128, 2017.
- [15] A. Rahmawati and E. Setyowati, "Analisis Cluster K-Means untuk Pengelompokan Kabupaten/Kota di Provinsi Bengkulu Berdasarkan Jumlah Base Transceiver Station dan Kekuatan Sinyal Telepon Seluler," *Jurnal of Computing Engineering, System and Science*, vol. 8, no. 1, pp. 184-196, 2023.
- [16] Y. I. Harnanto, A. Rusgiyono and T. Wuryandari, "Penerapan Analisis Kluster Metode Ward Terhadap Kabupaten/Kota di Jawa Tengah Berdasarkan Pengguna Alat Kontrasepsi," *Jurnal Gaussin*, vol. 6, no. 4, pp. 528-537, 2017.

- [17] A. A. Dzikrullah, "Pengelompokan Provinsi Berdasarkan Kualitas Jaringan Internet Dengan Metode Centroid Linkage," *Journal of Mathematics, Computations, and Statistics*, vol. 5, no. 1, pp. 48-57, April 2022.