

**ANALISIS CLUSTER DENGAN AVERAGE LINKAGE METHOD
DAN WARD’S METHOD PADA PENGELOMPOKAN
KABUPATEN/KOTA DI PROVINSI SUMATERA UTARA
BERDASARKAN INDIKATOR INDEKS PEMBANGUNAN
MANUSIA TAHUN 2022**

Putri Nazwa Maharani¹, Sawaluddin²
Universitas Sumatera Utara
e-mail: pnazwa340@gmail.com¹, sawal@usu.ac.id²

INFORMASI ARTIKEL

Submitted	: 2024-12-31
Review	: 2024-12-31
Accepted	: 2024-12-31
Published	: 2024-12-31

KATA KUNCI

Analisis Cluster, Average Linkage Method, Ward’s Method, Indeks Pembangunan Manusia, Sumatera Utara.

A B S T R A K

Analisis cluster merupakan salah satu metode interdependensi yang dapat digunakan untuk menentukan seberapa dekat atau mirip objek dan variabel. Analisis cluster terbagi atas dua metode, yaitu hirarki dan nonhirarki. Dalam penelitian ini menggunakan analisis cluster hirarki dengan metode average linkage dan metode ward. Kedua metode ini diterapkan dalam pengelompokan kabupaten/kota di Provinsi Sumatera Utara berdasarkan indikator indeks pembangunan manusia. Dari 33 kabupaten/kota di Provinsi Sumatera Utara diperoleh hasil metode average linkage dan metode ward sebanyak 3 kelompok, yaitu kelompok tingkat rendah, kelompok tingkat sedang, dan kelompok tingkat tinggi. Kelompok yang terbentuk di antaranya, kelompok tingkat rendah terdiri dari 9 kabupaten/kota yaitu Pakpak Bharat, Padang Lawas, Nias Selatan, Nias Barat, Nias Utara, Humbang Hasundutan, Nias, Kota Gunungsitoli, Samosir. Kelompok tingkat sedang terdiri dari 23 kabupaten/kota yaitu Mandailing Natal, Kota Tanjungbalai, Tapanuli Selatan, Padang Lawas Utara, Batu Bara, Tapanuli Tengah, Asahan, Labuhanbatu Selatan, Tapanuli Utara, Serdang Bedagai, Langkat, Dairi, Kota Sibolga, Kota Padangsidimpuan, Labuhanbatu Utara, Labuhan Batu, Toba, Kota Tebing Tinggi, Simalungun, Karo, Deli Serdang, Kota Binjai, Kota Pematangsiantar. Kelompok tingkat tinggi terdiri dari satu anggota yaitu Kota Medan. Pada penelitian ini ward’s method terbukti lebih unggul karena memiliki nilai Sum of Squared Errors (SSE) yang lebih rendah, menunjukkan bahwa metode ini mampu membentuk cluster yang lebih padat dan homogen. Selain itu, evaluasi dendrogram menunjukkan bahwa ward’s method menghasilkan cluster yang lebih seragam dan konsisten.

A B S T R A C T

Cluster analysis is an interdependence technique that can describe the proximity or similarity between objects and variables. Cluster analysis is divided into two

Keywords: Cluster Analysis, Average Linkage Method, Ward’s Method, Human

Development Index, North Sumatra.

methods: hierarchical and non-hierarchical. This study employs hierarchical cluster analysis using the average linkage method and Ward's method. Both methods are applied to cluster the districts/cities in North Sumatra Province based on Human Development Index (HDI) indicators. From the 33 districts/cities in North Sumatra Province, both the average linkage and Ward's methods produced 3 groups: low-level, medium-level, and high-level clusters. The low-level cluster includes 9 districts/cities: Pakpak Bharat, Padang Lawas, Nias Selatan, Nias Barat, Nias Utara, Humbang Hasundutan, Nias, Gunungsitoli City, and Samosir. The medium-level cluster consists of 23 districts/cities: Mandailing Natal, Tanjungbalai City, South Tapanuli, Padang Lawas Utara, Batu Bara, Central Tapanuli, Asahan, South Labuhanbatu, North Tapanuli, Serdang Bedagai, Langkat, Dairi, Sibolga City, Padangsidimpuan City, North Labuhanbatu, Labuhan Batu, Toba, Tebing Tinggi City, Simalungun, Karo, Deli Serdang, Binjai City, and Pematangsiantar City. The high-level cluster consists of a single member, Medan City. In this study, Ward's method proved to be more effective as it achieved a lower Sum of Squared Errors (SSE), indicating that it can form more compact and homogeneous clusters. Additionally, dendrogram evaluation revealed that Ward's method produces clusters that are more uniform and consistent.

PENDAHULUAN

Sumatera Utara merupakan sebuah provinsi yang terletak di bagian utara Pulau Sumatera, Indonesia. Provinsi Sumatera Utara memiliki 33 kabupaten/kota, yang terdiri dari 25 kabupaten, 8 kota, 325 kecamatan, dan 5.456 kelurahan/desa. Provinsi ini memiliki populasi yang besar dan luas wilayah yang cukup signifikan. Sumatera Utara memiliki karakteristik geografis yang beragam, mulai dari dataran rendah di sepanjang pantai hingga pegunungan yang menjulang tinggi di pedalaman.

Sumatera Utara memiliki latar belakang yang mencerminkan dinamika pembangunan manusia di wilayah tersebut. Provinsi ini, meskipun memiliki potensi ekonomi dan sumber daya alam yang besar, masih dihadapkan pada sejumlah tantangan dalam meningkatkan indeks pembangunan manusia (IPM) penduduknya.

IPM Sumatera Utara 2022 adalah sebesar 72,71 atau tumbuh 0,99 persen (meningkat 0,71 poin) dibandingkan capaian IPM 2021. Namun salah satu aspek utama yang memengaruhi IPM Sumatera Utara adalah ketimpangan akses terhadap layanan kesehatan dan pendidikan. Di beberapa daerah, terutama di pedalaman dan daerah terpencil, akses terhadap fasilitas kesehatan dan pendidikan masih terbatas. Hal ini dapat memengaruhi kesehatan dan tingkat pendidikan masyarakat setempat, serta berdampak pada IPM secara keseluruhan (Badan Pusat Statistik Provinsi Sumatera Utara, 2022).

Selain itu, tingkat kemiskinan dan pengangguran juga merupakan faktor yang memengaruhi pembangunan manusia di Sumatera Utara. Meskipun terdapat potensi ekonomi yang besar, namun masih terdapat sejumlah penduduk yang hidup di bawah garis kemiskinan dan menghadapi kesulitan dalam memperoleh pekerjaan yang layak.

Ini dapat mempengaruhi kemampuan mereka untuk mengakses layanan kesehatan dan pendidikan yang memadai, serta memengaruhi standar hidup mereka secara keseluruhan.

Indeks Pembangunan Manusia (IPM) di Sumatera Utara menjadi sebuah kunci bagi pemerintah dalam mengarahkan pembangunan wilayahnya. IPM tidak sekadar mencerminkan tingkat pembangunan ekonomi, tetapi juga menggambarkan kualitas hidup penduduknya, termasuk aspek kesehatan, pendidikan, dan standar hidup.

Analisis cluster digunakan untuk mengidentifikasi pola-pola atau kelompok-kelompok yang ada dalam data IPM. Dengan cara ini, kita dapat melihat apakah ada daerah-daerah yang memiliki karakteristik pembangunan manusia yang serupa atau berbeda secara signifikan.

Analisis cluster merupakan teknik statistik multivariat yang berguna untuk mengelompokkan objek berdasarkan kemiripan variabel yang akan diamati. Tujuannya adalah membentuk kelompok dimana objek dalam kelompok memiliki kesamaan tinggi. Analisis cluster terbagi menjadi dua metode, yaitu metode hirarki dan non-hirarki. Kelebihan dari menggunakan metode hirarki yaitu mempercepat pengolahan serta meminimalkan waktu karena membentuk dendrogram atau tingkatan sendiri, memudahkan penafsiran (Johnson, 2009).

Metode K-means (C-means) merupakan analisis cluster jenis metode non-hirarki, sementara pada metode hirarki terdapat beberapa metode antara lain yaitu metode pautan tunggal (single linkage), pautan lengkap (complete linkage), pautan rata-rata (average linkage), dan metode ward. Average linkage yang melibatkan pengelompokan berdasarkan jarak rata-rata antar objeknya, sementara pada metode ward untuk menghitung jarak antar cluster biasanya menggunakan pendekatan analisis varians. Untuk mengukur jarak kemiripan dalam metode average linkage dan ward menggunakan jarak Euclidean. Jarak Euclidean digunakan untuk mencari kedekatan jarak dari dua variabel.

Penelitian sebelumnya, seperti yang dilakukan (Laeli, 2014) menunjukkan bahwa dengan menggunakan metode average linkage mempunyai kinerja yang lebih baik dibandingkan metode ward karena memiliki nilai rasio lebih kecil. Begitu pula (Nafisah & Chandra, 2017) yang dimana penelitiannya mendapatkan bahwa metode average linkage membentuk 3 kelompok yaitu kelompok tingkat rendah, kelompok tingkat sedang dan kelompok tingkat tinggi. Selain itu (Pratama & Nasrudin, 2022) melakukan penelitian yang mendapatkan hasil bahwa metode terbaik adalah ward dengan hasil empat klaster.

Untuk mewujudkan pemerataan tingkat ekonomi di berbagai wilayah Sumatera Utara, perlu dilakukan peningkatan yang lebih baik pada indikator Indeks Pembangunan Manusia (IPM). IPM adalah suatu metrik perbandingan yang mempertimbangkan harapan hidup, akses pendidikan, dan standar hidup dalam evaluasi pembangunan suatu wilayah atau negara. IPM digunakan untuk mengklasifikasikan tingkat kemajuan ekonomi, membedakan antara negara maju, berkembang, atau tertinggal, serta mengukur dampak kebijakan ekonomi terhadap kualitas hidup.

Sebagai ukuran kemajuan atau ketertinggalan suatu daerah, IPM menjadi tolok ukur signifikan. Tingkat IPM yang tinggi menandakan keberhasilan suatu daerah dalam pembangunan, sementara upaya peningkatan IPM membutuhkan dukungan pemerintah dalam merancang program dan fasilitas yang memadai. Dukungan dari masyarakat juga diperlukan agar program-program tersebut dapat diimplementasikan secara efektif. Dengan demikian, penting untuk mempertimbangkan pengelompokan karakteristik

kabupaten/kota di Provinsi Sumatera Utara berdasarkan Indeks Pembangunan Manusia (IPM), dalam bentuk skripsi atau tugas akhir dengan judul “Analisis Cluster dengan Average Linkage Method dan Ward’s Method pada Pengelompokan Kabupaten/Kota di Provinsi Sumatera Utara Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2022”.

METODE PENELITIAN

Pada penelitian ini akan menggunakan data sekunder dari website BPS mengenai indikator IPM pada tahun 2022. IPM menggunakan beberapa indikator utama yang terdiri dari variabel Angka Harapan Hidup saat Lahir, variabel Harapan Lama Sekolah, variabel Rata-rata Lama Sekolah, dan variabel Pengeluaran per Kapita.

Inisialisasi tiap variabel

Angka Harapan Hidup = AHH

Harapan Lama Sekolah = HLS

Rata-rata Lama Sekolah = RLS

Pengeluaran per Kapita = PPE

Tabel 1. Data kabupaten/kota setiap variabel

No/Titik	Kabupaten Kota	AHH	HLS	RLS	PPE
1	Nias	70.06	13.04	5.88	7196
2	Mandailing Natal	63.05	13.85	8.76	10061
3	Tapanuli Selatan	65.28	13.37	9.34	11727
4	Tapanuli Tengah	67.58	13.24	8.86	10495
5	Tapanuli Utara	69.16	13.72	10	11892
6	Toba	70.76	13.58	10.58	12475
7	Labuhan Batu	70.37	12.95	9.4	11474
8	Asahan	68.73	12.63	8.82	11515
9	Simalungun	71.78	12.81	9.63	11524
10	Dairi	69.64	13.24	9.72	10740
11	Karo	72.03	12.95	10.02	12554
12	Deli Serdang	72.07	13.38	10.27	12657
13	Langkat	69.39	13.26	8.68	11439
14	Nias Selatan	69.21	12.48	6.23	7195
15	Humbang Hasundutan	70.02	13.31	10	8250
16	Pakpak Bharat	66.44	13.89	9.39	8558
17	Samosir	71.82	13.5	9.46	8991
18	Serdang Bedagai	69.21	12.63	8.71	11499
19	Batu Bara	67.55	12.93	8.26	10755
No/Titik	Kabupaten Kota	AHH	HLS	RLS	PPE
20	Padang Lawas Utara	67.53	13.23	9.46	10447
21	Padang Lawas	67.43	13.41	9.31	9207
22	Labuhanbatu Selatan	69.16	13.12	8.92	11751
23	Labuanbatu Utara	69.91	13.29	8.64	12258
24	Nias Utara	69.9	13.06	6.78	6575
25	Nias Barat	69.43	12.97	6.97	6152
26	Kota Sibolga	69.77	13.3	10.43	12135
27	Kota Tanjungbalai	63.87	12.9	9.55	11589
28	Kota Pematangsiantar	74.25	14.59	11.31	12744
29	Kota Tebing Tinggi	71.29	12.91	10.65	13144
30	Kota Medan	73.58	14.77	11.5	15503
31	Kota Binjai	72.79	13.87	11.18	11358
32	Kota Padangsidimpuan	69.84	14.58	11.11	11337
33	Kota Gunungsitoli	71.71	13.77	8.64	8434

Metode Pegumpulan Data

Dalam penelitian ini untuk menganalisis cluster yaitu dengan menggunakan perbandingan antara average linkage dan ward's method.

Beberapa tahapan yang digunakan dalam pemecahan masalah pada penelitian ini yaitu:

1. Melakukan analisis statistika deskriptif pada kajian karakteristik untuk melihat bagaimana gambaran umum dari setiap variabel pada kabupaten/kota yang berada di Provinsi Sumatera Utara.
2. Melakukan uji asumsi pada data yaitu dengan uji multikolinearitas menggunakan statistik Variance Inflation Factor (VIF).
3. Memperoleh hasil kelompok kabupaten/kota di Provinsi Sumatera Utara berdasarkan indikator IPM tahun 2022 dengan metode cluster hierarki.
4. Metode cluster hierarki yang digunakan adalah metode average linkage, dan ward. Menentukan metode terbaik di antara 2 metode tersebut menggunakan nilai korelasi cophenetic.
5. Menentukan kelompok terbaik dari metode terbaik.

HASIL DAN PEMBAHASAN

1. Analisis Deskriptif

Pada analisis deskriptif ini peneliti akan melihat gambaran umum dari data yang digunakan yaitu indikator indeks pembangunan manusia pada Tabel 2.

Tabel 2 Deskriptif data

	AHH	HLS	RLS	PPE
Mean	69.53	13.35	9.29	10716.09
Quartil 1	68.16	12.95	8.70	9099
Median	69.77	13.26	9.40	11439
Quartil 3	71.50	13.65	10.15	12013.50
Minimum	63.05	12.48	5.88	6152
Maximum	74.25	14.77	11.50	15503

Berdasarkan Tabel 2 variabel angka harapan hidup memiliki nilai minimum 63.05, nilai maksimum 74.25, dan nilai rata-rata sebesar 69.53. Nilai median, kuartil 1, dan kuartil 3 berturut-turut sebesar 69.77, 68.16, dan 71.50. Variabel harapan lama sekolah memiliki nilai minimum 12.48, nilai maksimum 14.77, dan nilai rata-rata sebesar 13.35 dengan sebaran nilai median, kuartil 1, dan kuartil 3 sebesar 13.26, 12.95, dan 13.65. Variabel rata-rata lama sekolah memiliki nilai rata-rata sebesar 9.29 dengan nilai minimum 5.88 dan nilai maksimum 11.50. Nilai median, kuartil 1, dan kuartil 3 sebesar 9.40, 8.70, dan 10.15. Variabel pengeluaran per kapita memiliki nilai median sebesar 11439 dengan nilai kuartil 1 sebesar 9099 dan nilai kuartil 3 sebesar 12013.50. Nilai minimum dan nilai maksimumnya sebesar 6152 dan 15503. Nilai rata-ratanya sebesar 10716.1.

2. Uji Multikolinearitas dengan VIF

Analisis cluster hierarki diawali dengan uji asumsi analisis cluster. Asumsi analisis cluster hierarki yang pertama yaitu pengujian multikolinearitas menggunakan nilai VIF (Variance Inflation Factor). Apabila nilai VIF berada dibawah 10, maka disimpulkan tidak terjadi multikolinearitas.

Tabel 3. Uji multikolinearitas

	VIF
AHH	1.132852
HLS	1.704182
RLS	3.716231
PPE	2.587964

Berdasarkan Tabel 3. dapat diketahui nilai variance inflation factor (VIF) keempat variabel memiliki nilai variance inflation factor (VIF) lebih kecil dari 10, sehingga bisa disimpulkan bahwa antar variabel tidak terjadi multikolinearitas.

3. Matriks Jarak *Euclidean Kuadrat*

Mencari jarak *Euclidean* kuadrat menggunakan rumus (2.1)

Pada Tabel 1, misalkan variabel:

Angka Harapan Hidup = x

Harapan Lama Sekolah = y

Rata-rata Lama Sekolah = z

Pengeluaran Per Kapita = w

$$\begin{aligned}
 d^2(1,2) &= (70.06 - 63.05)^2 + (13.04 - 13.85)^2 + (5.88 - 8.76)^2 \\
 &\quad + (7196 - 10061)^2 \\
 &= (7.01)^2 + (-0.81)^2 + (-2.88)^2 + (-2865)^2 \\
 &= 49.1401 + 0.6561 + 8.2944 + 8208225 \\
 &= 8208283.091
 \end{aligned}$$

$$\begin{aligned}
 d^2(1,3) &= (70.06 - 65.28)^2 + (13.04 - 13.37)^2 + (5.88 - 9.34)^2 \\
 &\quad + (7196 - 11727)^2 \\
 &= (4.78)^2 + (-0.33)^2 + (-3.46)^2 + (-4531)^2 \\
 &= 22.8484 + 0.1089 + 11.9716 + 20529961 \\
 &= 20529995.93
 \end{aligned}$$

$$\begin{aligned}
 d^2(1,4) &= (70.06 - 67.58)^2 + (13.04 - 13.24)^2 + (5.88 - 8.86)^2 \\
 &\quad + (7196 - 10495)^2 \\
 &= (2.48)^2 + (-0.2)^2 + (-2.98)^2 + (-3299)^2 \\
 &= 6.1504 + 0.04 + 8.8804 + 10883401 \\
 &= 10883416.07
 \end{aligned}$$

⋮

$$\begin{aligned}
 d^2(32,33) &= (69.84 - 71.71)^2 + (14.58 - 13.77)^2 + (11.11 - 8.64)^2 \\
 &\quad + (11337 - 8434)^2 \\
 &= 8427419.254
 \end{aligned}$$

Mencari matriks jarak euclidean kuadrat dilakukan hingga semua titik memiliki nilai matriks jarak.

Tabel 4. Matriks jarak euclidean kuadrat

	1: Nias	2: Mandailing Natal	3: Tapanuli Selatan	4: Gunung Sitoli
1: Nias	0.00	8208283.09	20529995.93	1532654.87
2: Mandailing Natal	8208283.09	0.00	2775561.54	2647204.02
3: Tapanuli Selatan	20529995.93	2775561.54	0.00	10843890.99

4: Tapanuli Tengah	10883416.07	188376.90	1517829.54	.	4247738.39
5: Tapanuli Utara	22052434.25	3352599.89	27240.61	.	11957772.35
6: Toba	27867863.87	5827458.83	559535.61	.	16329685.70
7: Labuhan Batu	18301296.49	1996623.80	64035.09	.	9241603.05
8: Asahan	18653771.58	2114149.75	44956.72	.	9492571.21
9: Simalungun	18731601.07	2140447.05	41251.65	.	9548101.91
10: Dairi	12559950.96	461085.72	974188.17	.	5317641.73
11: Karo	28708185.03	6215132.04	683975.20	.	16974402.68
12: Deli Serdang	29822544.43	6739299.86	864946.97	.	17833731.94
13: Langkat	18003057.34	1898924.55	82961.34	.	9030030.64
14: Nias Selatan	2.16	8214002.22	20539049.91	.	1535134.72
15: Humbang					
Hasundutan	1110933.05	3279771.41	12089551.91	.	33860.92
16: Pakpak Bharat	1855070.15	2259020.89	10042562.62	.	15404.35
17: Samosir	3222041.13	1144977.53	7485738.80	.	310249.76
18: Serdang Bedagai	18515817.90	2067883.44	52000.39	.	9394232.55
19: Batu Bara	12666492.98	481657.35	944790.51	.	5387059.16
32: Padangsidimpuan	17147910.77	1628228.16	152125.39	.	8427419.25
33: Gunungsitoli	1532654.87	2647204.02	10843890.99	.	0.00

4. Analisis Cluster

1) Average Linkage

Menggabungkan dua titik cluster yang memiliki jarak terdekat

Titik 1 = Cluster 1

Titik 2 = Cluster 2

Titik 3 = Cluster 3

⋮

Titik 33 = Cluster 33

Pada Tabel 4, jarak terdekat adalah 2.159 terletak pada titik 1,14. Maka,

Cluster A = {1,14}

$$Centroid_A = \left(\frac{70.06 + 63.05}{2}, \frac{13.04 + 12.48}{2}, \frac{5.88 + 6.23}{2}, \frac{7196 + 7195}{2} \right)$$

$$= \left(\frac{139.27}{2}, \frac{25.52}{2}, \frac{12.11}{2}, \frac{14391}{2} \right)$$

$$= (69.64, 12.76, 6.06, 7195.5)$$

SSE mengukur seberapa baik centroid mewakili cluster. Semakin kecil SSE, semakin baik representasi centroid.

$$SSE_A = (70.06 - 69.64)^2 + (13.04 - 12.76)^2 + (5.88 - 6.06)^2 + (7196 - 7195.5)^2$$

$$+ (63.05 - 69.64)^2 + (12.48 - 12.76)^2 + (6.23 - 6.06)^2$$

$$+ (7195 - 7195.5)^2$$

$$= 0.180625 + 0.0784 + 0.0324 + 0.25 + 0.180625 + 0.0729 + 0.0289$$

$$+ 0.25$$

$$= 0.541425 + 0.532425 = 1.07385$$

Memperbarui matriks jarak dengan average linkage dengan rumus (2.2)

$$D(A, 2) = \frac{d(1,2) + d(14,2)}{2 \times 1} = \frac{8208283.09 + 8214002.22}{2} = 8211142.66$$

$$D(A, 3) = \frac{d(1,3) + d(14,3)}{2 \times 1} = \frac{20529995.93 + 20539049.91}{2} = 20534522.92$$

$$D(A, 4) = \frac{d(1,4) + d(14,4)}{2 \times 1} = \frac{10883416.07 + 10890010.15}{2} = 10886713.11$$

$$D(A, 5) = \frac{d(1,5) + d(14,5)}{2 \times 1} = \frac{22052434.25 + 22061824.75}{2} = 22057129.50$$

$$D(A, 33) = \frac{d(1,33) + d(14,33)}{2 \times 1} = \frac{1532654.87 + 1535134.72}{2} = 1533894.80$$

Tabel 5. Matriks baru cluster A

	A(1,14)	2	3	4	.	33
A(1,14)	0	8211142.66	20534522.92	10886713.11	.	1533894.80
2		0.00	2775561.54	188376.90	.	2647204.02
3			0.00	1517829.54	.	10843890.99
4				0.00	.	4247738.39
5					.	11957772.35
.	.	.	.	.	.	.
33					.	0

Pada Tabel 5. jarak terdekat adalah 90.991 pada titik terletak pada titik 8,9. Maka, Cluster B = {8,9}.

$$Centroid_B = \left(\frac{68.73 + 71.78}{2}, \frac{12.63 + 12.81}{2}, \frac{8.82 + 9.63}{2}, \frac{11515 + 11524}{2} \right)$$

$$= (70.26, 12.72, 9.23, 11519.5)$$

$$SSE_B = (68.73 - 70.26)^2 + (12.63 - 12.72)^2 + (8.82 - 9.23)^2 + (11515 - 11519.5)^2$$

$$+ (71.78 - 70.26)^2 + (12.81 - 12.72)^2 + (9.63 - 9.23)^2$$

$$+ (11524 - 11519.5)^2$$

$$= 2.33 + 0.01 + 0.16 + 20.25 + 2.33 + 0.01 + 0.16 + 20.25$$

$$= 22.75 + 22.75 = 45.50$$

Memperbarui matriks jarak

$$D(B, A) = \frac{d(8, A) + d(9, A)}{2 \times 2} = \frac{18653771.58 + 18731601.074}{4} = 9346343.164$$

$$D(B, 2) = \frac{d(8,2) + d(9,2)}{2 \times 1} = \frac{2114149.75 + 2140447.05}{2} = 2127298.403$$

$$D(B, 3) = \frac{d(8,3) + d(9,3)}{2 \times 1} = \frac{44956.72 + 41251.6477}{2} = 43104.184$$

$$D(B, 4) = \frac{d(8,4) + d(9,4)}{2 \times 1} = \frac{1040401.70 + 1058859.42}{2} = 1049630.557$$

$$D(B, 33) = \frac{d(8,33) + d(9,33)}{2 \times 1} = \frac{9492571.21 + 9548101.90}{2} = 9520336.56$$

Tabel 6. Matriks baru cluster B

	A(1,14)	B(8,9)	2	3	.	33
A(1,14)	0	9346343.16	8211142.66	20534522.92	.	1533894.80
B(8,9)		0.00	2127298.40	43104.18	.	9520336.56
2			0.00	2775561.54	.	2647204.02
3				0.00	.	10843890.99
4					.	4247738.39
.	.	.	.	.	.	.
33					.	0.00

Pada Tabel 6. jarak terdekat adalah 231.596 terletak pada titik 10,19. Maka, Cluster C = {10,19}.

$$Centroid_C = \left(\frac{69.64 + 67.55}{2}, \frac{13.24 + 12.93}{2}, \frac{9.72 + 8.26}{2}, \frac{10740 + 10755}{2} \right)$$

$$= (68.60, 13.09, 8.99, 10747.5)$$

$$SSE_C = (69.64 - 68.60)^2 + (13.24 - 13.09)^2 + (9.72 - 8.99)^2 + (10740 - 10747.50)^2$$

$$+ (67.55 - 68.60)^2 + (12.93 - 13.09)^2 + (8.26 - 8.99)^2$$

$$+ (10755 - 10747.50)^2$$

$$= 1.09 + 0.02 + 0.53 + 56.25 + 1.09 + 0.02 + 0.53 + 56.25$$

$$= 57.90 + 57.90 = 115.80$$

Memperbarui matriks jarak

$$D(C, A) = \frac{d(10, A) + d(19, A)}{2 \times 2} = \frac{12559950.962 + 12666492.9766}{4} = 6306610.99$$

$$D(C, B) = \frac{d(10, B) + d(19, B)}{2 \times 2} = \frac{607643.8914 + 584491.2901}{4} = 298033.80$$

$$D(C, 2) = \frac{d(10, 2) + d(19, 2)}{2 \times 1} = \frac{461085.7218 + 481657.3464}{2} = 471371.53$$

$$D(C, 3) = \frac{d(10, 3) + d(19, 3)}{2 \times 1} = \frac{974188.1709 + 944790.5129}{2} = 959489.342$$

$$D(C, 33) = \frac{d(10, 33) + d(19, 33)}{2 \times 1} = \frac{5317641.7322 + 5387059.1556}{2} = 5352350.444$$

Tabel 7. Matriks baru cluster C

	A(1,14)	B (8,9)	C (10,19)	2	.	33
A (1,14)	0	9346343.16	6306610.98	8211142.66	.	1533894.80
B (8,9)		0.00	298033.80	2127298.40	.	9520336.56
C(10,19)			0.00	471371.53	.	5352350.44
2				0.00	.	2647204.02
3					.	10843890.99
.	.	.	.	.	.	.
33					.	0

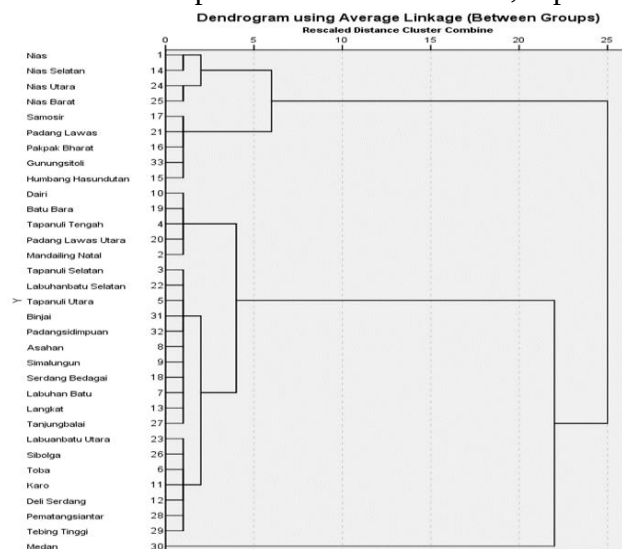
Pada Tabel 7. jarak terdekat adalah 444.363 terletak pada titik B,18.

Memperbarui matriks jarak dan mencari jarak terdekat sampai membentuk 3 cluster akhir.

Tabel 8. Hasil jarak terdekat semua cluster

Cluster 1	Cluster 2	Nilai Terkecil
1	14	2.159
8	9	90.991
10	19	231.596
8	18	444.363
31	32	450.212
3	22	591.293
7	13	1226.575
4	20	2304.363
7	8	3570.369
6	11	6243.323
12	28	7576.298
7	27	10743.147
23	26	15132.224
16	33	15404.350
3	5	23561.569
7	31	27606.812
6	12	38054.779
17	21	46675.303
15	16	64369.221
4	10	77087.373
3	7	116105.930
24	25	178929.265
6	23	183088.547
Cluster 1	Cluster 2	Nilai Terkecil
2	4	320034.229
15	17	496909.789
6	29	499321.712
1	24	736957.594
3	6	1145107.220
2	3	2459276.829
1	15	3964694.818

Penggabungan dilakukan sampai membentuk 3 cluster, seperti dendrogram berikut:



Gambar 1. Dendrogram average linkage

Pada Gambar 1, sumbu horizontal menunjukkan jarak antar cluster. Semakin jauh suatu cabang menyebar ke kanan, semakin besar jarak antar cluster yang digabungkan pada titik tersebut. Dengan kata lain, jarak ini menunjukkan tingkat perbedaan antara kabupaten/kota atau antar cluster yang digabungkan. Pada jarak yang lebih tinggi terlihat tiga cluster besar yang terbentuk, mencerminkan kelompok kabupaten/kota dengan tingkat IPM rendah yang ditandai dengan sumbu kuning, sedang yang ditandai dengan sumbu merah, dan tinggi yang ditandai dengan sumbu biru.

Data diurutkan berdasarkan hasil cluster

Tabel 9. Hasil cluster

Kabupaten/Kota	AHH	HLS	RLS	PPE	Cluster
Nias	70.06	13.04	5.88	7196	1
Nias Selatan	69.21	12.48	6.23	7195	1
Humbang Hasundutan	70.02	13.31	10	8250	1
Pakpak Bharat	66.44	13.89	9.39	8558	1
Samosir	71.82	13.5	9.46	8991	1
Padang Lawas	67.43	13.41	9.31	9207	1
Nias Utara	69.9	13.06	6.78	6575	1
Nias Barat	69.43	12.97	6.97	6152	1
Kota Gunungsitoli	71.71	13.77	8.64	8434	1
Mandailing Natal	63.05	13.85	8.76	10061	2
Tapanuli Selatan	65.28	13.37	9.34	11727	2
Tapanuli Tengah	67.58	13.24	8.86	10495	2
Tapanuli Utara	69.16	13.72	10	11892	2
Toba	70.76	13.58	10.58	12475	2
Labuhan Batu	70.37	12.95	9.4	11474	2
Asahan	68.73	12.63	8.82	11515	2
Simalungun	71.78	12.81	9.63	11524	2
Dairi	69.64	13.24	9.72	10740	2
Karo	72.03	12.95	10.02	12554	2
Deli Serdang	72.07	13.38	10.27	12657	2
Langkat	69.39	13.26	8.68	11439	2
Serdang Bedagai	69.21	12.63	8.71	11499	2
Batu Bara	67.55	12.93	8.26	10755	2
Padang Lawas Utara	67.53	13.23	9.46	10447	2
Labuhanbatu Selatan	69.16	13.12	8.92	11751	2
Labuanbatu Utara	69.91	13.29	8.64	12258	2
Kota Sibolga	69.77	13.3	10.43	12135	2
Kota Tanjungbalai	63.87	12.9	9.55	11589	2
Kabupaten/Kota	AHH	HLS	RLS	PPE	Cluster
Kota Pematangsiantar	74.25	14.59	11.31	12744	2
Kota Tebing Tinggi	71.29	12.91	10.65	13144	2
Kota Binjai	72.79	13.87	11.18	11358	2
Kota Padangsidimpuan	69.84	14.58	11.11	11337	2
Kota Medan	73.58	14.77	11.5	15503	3

Nama masing-masing cluster berdasarkan karakteristik data yang dimiliki

Cluster 1 (Rendah): Cluster dengan nilai rata-rata terendah.

Cluster 2 (Sedang): Cluster dengan nilai rata-rata menengah.

Cluster 3 (Tinggi): Cluster dengan nilai rata-rata tertinggi.

Menghitung rata-rata dari setiap variabel dalam masing masing cluster

Tabel 10. Hasil rata-rata tiap cluster dengan average linkage

	Cluster 1 (Rendah)	Cluster 2 (Sedang)	Cluster 3 (Tinggi)
AHH	69.56	69.35	73.58
HLS	13.27	13.32	14.77
RLS	8.07	9.67	11.5
PPE	7839.78	11633.48	15503
Rata-rata	1982.67	2931.45	3900.71

Dari Gambar 1. dan Tabel 9. metode average linkage menghasilkan 3 kelompok cluster yaitu:

1. Cluster 1 memiliki anggota sebanyak 9 kabupaten/kota yaitu Nias, Nias Selatan, Humbang Hasundutan, Pakpak Bharat, Samosir, Padang Lawas, Nias Utara, Nias Barat, Kota Gunungsitoli yang memiliki indikator indeks pembangunan manusia dalam kategori rendah dengan nilai rata-rata sebesar 1982.67.
2. Cluster 2 memiliki anggota sebanyak 23 kabupaten/kota yaitu Mandailing Natal, Tapanuli Selatan, Tapanuli Tengah, Tapanuli Utara, Toba, Labuhan Batu, Asahan, Simalungun, Dairi, Karo, Deli Serdang, Langkat, Serdang Bedagai, Batu Bara, Padang Lawas Utara, Labuhanbatu Selatan, Labuhanbatu Utara, Kota Sibolga, Kota Tanjungbalai, Kota Pematang Siantar, Kota Tebing Tinggi, Kota Binjai, Kota Padangsidimpuan yang memiliki indikator indeks pembangunan manusia dalam kategori sedang dengan nilai rata-rata sebesar 2931.45.
3. Cluster 3 memiliki satu anggota yaitu Kota Medan yang memiliki indikator indeks pembangunan manusia dalam kategori tinggi dengan nilai rata-rata 3900.71.

2) Ward's Method

Menghitung varians berdasarkan hasil matriks jarak euclidean kuadrat

Titik 1 = Cluster 1

Titik 2 = Cluster 2

Titik 3 = Cluster 3

⋮

Titik 33 = Cluster 33

Tabel 11. Data dengan nilai centroid

Cluster	Kabupaten Kota	Centroid			
		AHH	HLS	RLS	PPE
C1	Nias	70.06	13.04	5.88	7196
C2	Mandailing Natal	63.05	13.85	8.76	10061
C3	Tapanuli Selatan	65.28	13.37	9.34	11727
C4	Tapanuli Tengah	67.58	13.24	8.86	10495
C5	Tapanuli Utara	69.16	13.72	10	11892
C6	Toba	70.76	13.58	10.58	12475
C7	Labuhan Batu	70.37	12.95	9.4	11474
C8	Asahan	68.73	12.63	8.82	11515
C9	Simalungun	71.78	12.81	9.63	11524
C10	Dairi	69.64	13.24	9.72	10740
C11	Karo	72.03	12.95	10.02	12554
C12	Deli Serdang	72.07	13.38	10.27	12657
C13	Langkat	69.39	13.26	8.68	11439
C14	Nias Selatan	69.21	12.48	6.23	7195
C15	Humbang Hasundutan	70.02	13.31	10	8250
C16	Pakpak Bharat	66.44	13.89	9.39	8558
C17	Samosir	71.82	13.5	9.46	8991
C18	Serdang Bedagai	69.21	12.63	8.71	11499
C19	Batu Bara	67.55	12.93	8.26	10755
C20	Padang Lawas Utara	67.53	13.23	9.46	10447

Analisis Cluster Dengan Average Linkage Method Dan Ward's Method Pada Pengelompokan Kabupaten/Kota Di Provinsi Sumatera Utara Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2022.

C21	Padang Lawas	67.43	13.41	9.31	9207
C22	Labuhanbatu Selatan	69.16	13.12	8.92	11751
C23	Labuanbatu Utara	69.91	13.29	8.64	12258
Cluster	Kabupaten Kota	Centroid			
		AHH	HLS	RLS	PPE
C24	Nias Utara	69.9	13.06	6.78	6575
C25	Nias Barat	69.43	12.97	6.97	6152
C26	Kota Sibolga	69.77	13.3	10.43	12135
C27	Kota Tanjungbalai	63.87	12.9	9.55	11589
C28	Kota Pematangsiantar	74.25	14.59	11.31	12744
C29	Kota Tebing Tinggi	71.29	12.91	10.65	13144
C30	Kota Medan	73.58	14.77	11.5	15503
C31	Kota Binjai	72.79	13.87	11.18	11358
C32	Kota Padangsidimpuan	69.84	14.58	11.11	11337
C33	Kota Gunungsitoli	71.71	13.77	8.64	8434

Menghitung nilai varians dengan rumus (2.3).

$$D(1,2) = \frac{1 \times 1}{1 + 1} \times 8208283.091 = 4104141.546$$

$$D(1,3) = \frac{1 \times 1}{1 + 1} \times 20529995.93 = 10264997.964$$

$$D(1,4) = \frac{1 \times 1}{1 + 1} \times 10883416.07 = 5441708.035$$

$$D(1,5) = \frac{1 \times 1}{1 + 1} \times 22052434.25 = 11026217.123$$

⋮

$$D(32,33) = \frac{1 \times 1}{1 + 1} \times 8427419.254 = 766327.437$$

Tabel 12. Hasil menghitung varians

	1: Nias	2: Mandailing Natal	3: Tapanuli Selatan	.	33: Gunungsitoli
1: Nias	0.00	4104141.55	10264997.96	.	766327.44
2: Mandailing Natal	4104141.55	0.00	1387780.77	.	1323602.01
3: Tapanuli Selatan	10264997.96	1387780.77	0.00	.	5421945.50
4: Tapanuli Tengah	5441708.04	94188.45	758914.77	.	2123869.19
5: Tapanuli Utara	11026217.12	1676299.94	13620.31	.	5978886.18
6: Toba	13933931.94	2913729.41	279767.81	.	8164842.85
7: Labuhan Batu	9150648.25	998311.90	32017.54	.	4620801.52
8: Asahan	9326885.79	1057074.88	22478.36	.	4746285.61
9: Simalungun	9365800.54	1070223.53	20625.82	.	4774050.95
10: Dairi	6279975.48	230542.86	487094.09	.	2658820.87
11: Karo	14354092.51	3107566.02	341987.60	.	8487201.34
12: Deli Serdang	14911272.21	3369649.93	432473.48	.	8916865.97
13: Langkat	9001528.67	949462.28	41480.67	.	4515015.32
14: Nias Selatan	1.08	4107001.11	10269524.95	.	767567.36
15: Humbang Hasundutan	555466.52	1639885.71	6044775.95	.	16930.46
16: Pakpak Bharat	927535.07	1129510.45	5021281.31	.	7702.17
17: Samosir	1611020.56	572488.76	3742869.40	.	155124.88
18: Serdang Bedagai	9257908.95	1033941.72	26000.19	.	4697116.28
19: Batu Bara	6333246.49	240828.67	472395.26	.	2693529.58

32: Padangsidempuan	8573955.39	814114.08	76062.70	.	4213709.63
33: Gunungsitoli	766327.44	1323602.01	5421945.50	.	0.00

Pada Tabel 12. jarak terdekat adalah 1.079 terletak pada titik 1,4. Maka, Cluster A = {1,14}.

$$Centroid_A = \left(\frac{70.06 + 63.05}{2}, \frac{13.04 + 12.48}{2}, \frac{5.88 + 6.23}{2}, \frac{7196 + 7195}{2} \right)$$

$$= \left(\frac{139.27}{2}, \frac{25.52}{2}, \frac{12.11}{2}, \frac{14391}{2} \right)$$

$$= (69.64, 12.76, 6.06, 7195.5)$$

$$Centroid_2 = (63.05, 13.85, 8.76, 10061)$$

$$\begin{aligned} \|Centroid_A - Centroid_2\|^2 &= (69.64 - 63.05)^2 + (12.76 - 13.85)^2 + (6.06 - 8.76)^2 \\ &\quad + (7195.5 - 10061)^2 \end{aligned}$$

$$= 43.43 + 1.19 + 7.29 + 8211090.25 = 8211142.16$$

$$D(A, 2) = \frac{2 \times 1}{2 + 1} \times 8211142.16 = 5474094.77$$

$$\begin{aligned} \|Centroid_A - Centroid_3\|^2 &= (69.64 - 65.28)^2 + (12.76 - 13.37)^2 + (6.06 - 9.34)^2 \\ &\quad + (7195.5 - 11727)^2 \end{aligned}$$

$$= 19.01 + 0.37 + 10.76 + 20534492.25 = 20534522.39$$

$$D(A, 3) = \frac{2 \times 1}{2 + 1} \times 20534522.39 = 13689681.59$$

$$\begin{aligned} \|Centroid_A - Centroid_4\|^2 &= (69.64 - 67.58)^2 + (12.76 - 13.24)^2 + (6.06 - 8.86)^2 \\ &\quad + (7195.5 - 10495)^2 \end{aligned}$$

$$= 4.24 + 0.23 + 7.84 + 10886700.25 = 10886712.56$$

$$D(A, 5) = \frac{2 \times 1}{2 + 1} \times 10886712.56 = 7257808.37$$

$$\begin{aligned} \|Centroid_A - Centroid_5\|^2 &= (69.64 - 69.16)^2 + (12.76 - 13.72)^2 + (6.06 - 10)^2 \\ &\quad + (7195.5 - 11892)^2 \end{aligned}$$

$$= 0.23 + 0.92 + 15.52 + 22057112.25 = 22057128.92$$

$$D(A, 5) = \frac{2 \times 1}{2 + 1} \times 22057128.92 = 14704752.61$$

⋮

$$\begin{aligned} & \|Centroid_A - Centroid_{33}\|^2 \\ &= (69.64 - 71.71)^2 + (12.76 - 13.77)^2 + (6.06 - 8.64)^2 \\ &+ (7195.5 - 8434)^2 \\ &= 4.28 + 1.02 + 6.66 + 1533882.25 = 1533894.21 \end{aligned}$$

$$D(A, 5) = \frac{2 \times 1}{2 + 1} \times 1533894.21 = 1022596.14$$

Tabel 13. Matriks baru cluster A

	A(1,14)	2	3	4	.	33
A(1,14)	0	5474094.77	13689681.59	7257808.38	.	1022596.14
2		0.00	1387780.77	94188.45	.	1323602.01
3			0.00	758914.77	.	5421945.50
4				0.00	.	2123869.19
.	.	.	.	.	.	.
33					.	0.00

Pada Tabel 13. jarak terdekat adalah 45.4955 terletak pada titik 8,9. Maka, Cluster B = {8,9}.

$$\begin{aligned} Centroid_B &= \left(\frac{68.73 + 71.78}{2}, \frac{12.63 + 12.81}{2}, \frac{8.82 + 9.63}{2}, \frac{11515 + 11524}{2} \right) \\ &= (70.26, \quad 12.72, \quad 9.23, \quad 11519.5) \end{aligned}$$

$$Centroid_A = (69.64, \quad 12.76, \quad 6.06, \quad 7195.5)$$

$$\begin{aligned} & \|Centroid_B - Centroid_A\|^2 \\ &= (70.26 - 69.64)^2 + (12.72 - 12.76)^2 + (9.23 - 6.06)^2 \\ &+ (11519.5 - 7195.5)^2 = 18696986.43 \end{aligned}$$

$$D(B, A) = \frac{2 \times 2}{2 + 2} \times 18696986.43 = 18696986.43$$

$$\begin{aligned} & \|Centroid_B - Centroid_2\|^2 \\ &= (70.26 - 63.05)^2 + (12.72 - 13.85)^2 + (9.23 - 8.76)^2 \\ &+ (11519.5 - 10061)^2 = 2127275.73 \end{aligned}$$

$$D(B, 2) = \frac{2 \times 1}{2 + 1} \times 2127275.73 = 1418183.82$$

$$\begin{aligned} & \|Centroid_B - Centroid_3\|^2 \\ &= (70.26 - 65.28)^2 + (12.72 - 13.37)^2 + (9.23 - 9.34)^2 \\ &+ (11519.5 - 11727)^2 = 43081.49 \end{aligned}$$

$$D(B, 3) = \frac{2 \times 1}{2 + 1} \times 43081.49 = 28720.99$$

$$\begin{aligned} & \|Centroid_B - Centroid_4\|^2 \\ &= (70.26 - 67.58)^2 + (12.72 - 13.24)^2 + (9.23 - 8.86)^2 \\ &+ (11519.5 - 10495)^2 = 1049607.84 \end{aligned}$$

$$D(B, 3) = \frac{2 \times 1}{2 + 1} \times 43081.49 = 28720.99$$

$$\begin{aligned} & \|Centroid_B - Centroid_4\|^2 \\ &= (70.26 - 67.58)^2 + (12.72 - 13.24)^2 + (9.23 - 8.86)^2 \\ &+ (11519.5 - 10495)^2 = 1049607.84 \end{aligned}$$

$$D(B, 5) = \frac{2 \times 1}{2 + 1} \times 1049607.84 = 699738.56$$

$$\begin{aligned} & \|Centroid_B - Centroid_{33}\|^2 \\ &= (70.26 - 71.71)^2 + (12.72 - 13.77)^2 + (9.23 - 8.64)^2 \\ &+ (11519.5 - 8434)^2 = 9520313.80 \end{aligned}$$

$$D(B, 33) = \frac{2 \times 1}{2 + 1} \times 9520313.80 = 6346875.87$$

Tabel 14. Matriks baru cluster B

	A(1,14)	B(8,9)	2	3	.	33
A(1,14)	0	18696986.47	5474094.77	13689681.59	.	1022596.14
B(8,9)		0.00	1418183.82	28720.99	.	6346875.87
2			0.00	1387780.77	.	1323602.01
3				0.00	.	5421945.50
4					.	2123869.19
.	.	.	.	.	.	.
33					.	0.00

Pada Tabel 14. jarak terdekat adalah 115.798 terletak pada titik 10,19. Maka, Cluster C = {10,19}.

$$Centroid_C = \left(\frac{69.64 + 67.55}{2}, \frac{13.24 + 12.93}{2}, \frac{9.72 + 8.26}{2}, \frac{10740 + 10755}{2} \right)$$

$$= (68.60, \quad 13.09, \quad 8.99, \quad 10747.5)$$

$$Centroid_A = (69.64, \quad 13.09, \quad 6.06, \quad 7195.5)$$

$$\begin{aligned} & \|Centroid_C - Centroid_A\|^2 \\ &= (68.60 - 69.64)^2 + (13.09 - 12.76)^2 + (8.99 - 6.06)^2 \\ &+ (10747.5 - 7195.5)^2 = 12616713.78 \end{aligned}$$

$$D(C, A) = \frac{2 \times 2}{2 + 2} \times 12616713.78 = 12616713.78$$

$$\begin{aligned} & \|Centroid_C - Centroid_B\|^2 \\ &= (68.60 - 70.26)^2 + (13.09 - 12.72)^2 + (8.99 - 9.23)^2 \\ &+ (10747.5 - 11519.5)^2 = 595986.95 \end{aligned}$$

$$D(C, B) = \frac{2 \times 2}{2 + 2} \times 595986.95 = 595986.95$$

$$\begin{aligned} \|Centroid_C - Centroid_2\|^2 &= (68.60 - 63.05)^2 + (13.09 - 13.85)^2 + (8.99 - 8.76)^2 \\ &+ (10747.5 - 10061)^2 = 471313.68 \end{aligned}$$

$$D(C, 2) = \frac{2 \times 1}{2 + 1} \times 471313.68 = 314209.12$$

$$\begin{aligned} \|Centroid_C - Centroid_3\|^2 &= (68.60 - 65.28)^2 + (13.09 - 13.37)^2 + (8.99 - 9.34)^2 \\ &+ (10747.5 - 11727)^2 = 959431.47 \end{aligned}$$

$$D(C, 3) = \frac{2 \times 1}{2 + 1} \times 959431.47 = 639620.98$$

$$\begin{aligned} \|Centroid_C - Centroid_{33}\|^2 &= (68.60 - 71.71)^2 + (13.09 - 13.77)^2 + (8.99 - 8.64)^2 \\ &+ (10747.5 - 8434)^2 = 5352292.51 \end{aligned}$$

$$D(C, 33) = \frac{2 \times 1}{2 + 1} \times 5352292.51 = 3568195.01$$

Tabel 15. Matriks baru cluster C

	A(1,14)	B(8,9)	C(10,19)	2	.	33
A(1,14)	0.00	18696986.47	12616713.78	5474094.77	.	1022596.14
B(8,9)		0.00	595986.95	1418183.82	.	6346875.87
C(10,19)			0.00	314209.12	.	3568195.00
2				0.00	.	1323602.01
3					.	5421945.50
.	.	.	.	.	.	.
33					.	0.00

Pada Tabel 15. jarak terdekat adalah 225.106 terletak pada titik 31,32.

Memperbarui matriks jarak dan mencari jarak terdekat sampai membentuk 3 cluster akhir.

Tabel 16. Hasil jarak terdekat semua cluster

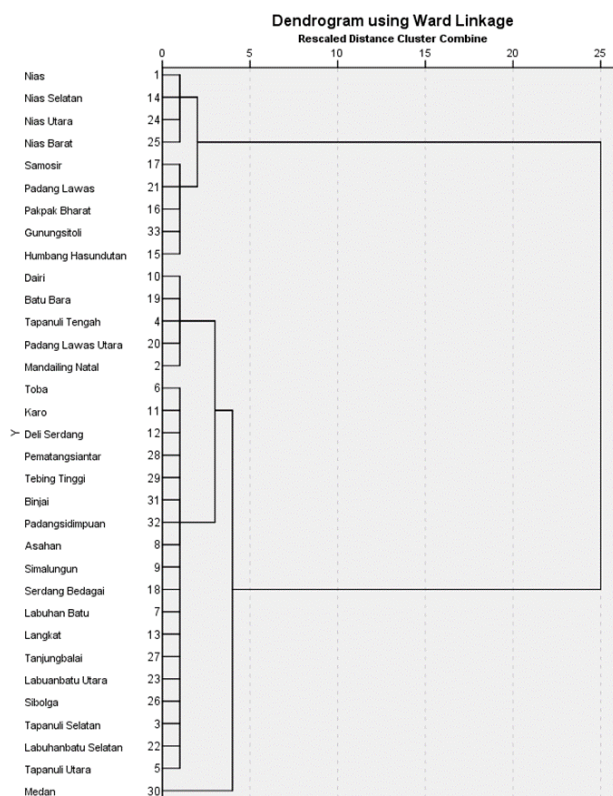
Cluster 1	Cluster 2	Nilai Terkecil
1	14	1.079
8	9	45.496
10	19	115.798
31	32	225.106
8	18	281087
3	22	295647
7	13	613.287
4	20	1152.181
6	11	3121.662
12	28	3788.149
23	26	20991.434
16	33	28693.609
7	27	36858.615
3	5	52467.779
17	21	75805.430

6	12	110405.304
7	31	148424.016
15	16	188769.438
4	10	265222.821
24	25	354687.454
3	23	552982.857
6	29	783249.855
2	4	1023732.965
15	17	1586803.083
1	24	2279027.821
3	7	3004964.857
3	6	7070927.985
1	15	15165108.280
2	3	23379210.264
2	30	37728547.890

Diperoleh hasil untuk ward method yaitu 3 kelompok cluster yang terdiri dari:

1. Cluster 1 memiliki anggota sebanyak 9 kabupaten/kota yaitu Pakpak Bharat, Padang Lawas, Nias Selatan, Nias Barat, Nias Utara, Humbang Hasundutan, Nias, Kota Gunungsitoli, Samosir.
2. Cluster 2 memiliki anggota sebanyak 23 kabupaten/kota yaitu Mandailing Natal, Kota Tanjungbalai, Tapanuli Selatan, Padang Lawas Utara, Batu Bara, Tapanuli Tengah, Asahan, Labuhanbatu Selatan, Tapanuli Utara, Serdang Bedagai, Langkat, Dairi, Kota Sibolga, Kota Padangsidempuan, Labuhanbatu Utara, Labuhan Batu, Toba, Kota Tebing Tinggi, Simalungun, Karo, Deli Serdang, Kota Binjai, Kota Pematangsiantar.
3. Cluster 3 memiliki satu anggota yaitu Kota Medan.

Penggabungan dilakukan sampai membentuk 3 cluster, seperti dendogram berikut:



Gambar 2. Dendrogram ward's method

Pada Gambar 2 sumbu horizontal menunjukkan jarak antar cluster. Semakin jauh suatu cabang menyebar ke kanan, semakin besar jarak antar cluster yang digabungkan pada titik tersebut. Dengan kata lain, jarak ini menunjukkan tingkat perbedaan antara kabupaten/kota atau antar cluster yang digabungkan. Pada jarak yang lebih tinggi terlihat tiga cluster besar yang terbentuk, mencerminkan kelompok kabupaten/kota dengan tingkat IPM rendah yang ditandai dengan sumbu kuning, sedang yang ditandai dengan sumbu merah, dan tinggi yang ditandai dengan sumbu biru.

3) Pemilihan Hasil Pengelompokkan Terbaik

Sebuah metode pengelompokkan dikatakan baik dengan menghitung jumlah kesalahan kuadrat (Sum of Squared Errors) dari setiap titik data terhadap centroid dalam cluster. Berdasarkan hasil perhitungan Sum of Squared Errors (SSE), ward's method menunjukkan nilai SSE yang lebih rendah dibandingkan average linkage. Nilai SSE yang lebih rendah pada ward's method menandakan bahwa kabupaten/kota dalam setiap cluster memiliki jarak yang lebih kecil satu sama lain, yang berarti bahwa anggota dalam cluster lebih mirip atau homogen.

Menentukan metode terbaik juga dapat dilakukan dengan mengevaluasi hasil dendrogram. Berdasarkan dendrogram, ward's method cenderung membentuk cluster yang lebih seragam. Ward's method menghasilkan cluster yang lebih kompak dengan nilai jarak yang lebih kecil, maka ward's method bisa dianggap sebagai metode yang lebih baik. Dengan demikian, ward's method memberikan keunggulan dalam kualitas cluster melalui nilai jarak yang lebih kecil.

KESIMPULAN

Dari hasil perhitungan dan pembahasan dapat ditarik kesimpulan:

1. Berdasarkan analisis cluster yang dilakukan dengan menggunakan average linkage method dan ward's method pada pengelompokan Kabupaten/Kota di Provinsi Sumatera Utara berdasarkan Indikator Indeks Pembangunan Manusia (IPM) Tahun 2022, kedua metode membentuk tiga cluster yang membedakan kabupaten/kota berdasarkan kesamaan dalam indikator IPM. Hasil perhitungan menunjukkan bahwa metode ward menghasilkan nilai Sum of Squared Errors (SSE) yang lebih rendah dibandingkan dengan average linkage. SSE yang lebih rendah ini mengindikasikan bahwa metode ward mampu membentuk cluster dengan kekompakan dan homogenitas yang lebih tinggi, sehingga kabupaten/kota dalam satu klaster memiliki kemiripan yang lebih kuat dalam indikator IPM. Selain itu, evaluasi dendrogram menunjukkan bahwa metode ward cenderung membentuk cluster yang lebih seragam dalam hal ukuran dan distribusi, yang berarti cluster yang dihasilkan lebih seimbang dan konsisten. Dengan demikian, metode ward menjadi metode terbaik dalam penelitian ini karena lebih efektif dalam menciptakan cluster yang terpusat dan seragam. Metode ward dapat dianggap lebih optimal untuk menggambarkan variasi tingkat kesejahteraan antar-kabupaten/kota di Sumatera Utara, yang ditunjukkan oleh kemiripan yang lebih jelas di dalam masing-masing cluster.
2. Berdasarkan hasil penelitian, cluster yang terbentuk adalah sebanyak 3 cluster dimana cluster 1 terdiri dari 9 anggota, cluster 2 terdiri dari 23 anggota, dan cluster 3 terdiri dari satu anggota saja.

Saran

Untuk penelitian selanjutnya dapat menggunakan metode cluster lain seperti K-Means, DBSCAN, atau metode hierarki lainnya untuk membandingkan hasil yang lebih komprehensif. Dengan menambahkan lebih banyak metode, analisis dapat

mengidentifikasi metode cluster yang paling efektif dan sesuai untuk data IPM, memberikan dasar yang lebih kuat dalam memilih teknik cluster terbaik.

DAFTAR PUSTAKA

- Ananto, S. T. (2023). Metode Perhitungan IPM (Indeks Pembangunan Manusia). Statistisi Ahli Madya. <https://www.slideshare.net/slideshow/metode-penghitungan-indeks-pembangunan-manusia-ipmpdfpdf/266949903#8>
- Anggeyeny, Viany Indah. (2009). “Fear of Floating: Studi Empiris Sistem Nilai Tukar secara de facto di Indonesia dalam Periode 1994-2003”. Universitas Indonesia.
- Badan Pusat Statistik Provinsi Sumatera Utara. (2022). Indeks Pembangunan Manusia (IPM) Sumatera Utara Tahun 2022.
- Badan Pusat Statistik. (2022). Indeks Pembangunan Manusia. <https://www.bps.go.id/id>
- Cahyoningtyas, R. A. (2019). Metode Ward dan Average Linkage Clustering untuk Segmentasi Objek Wisata di Malang Raya. FMIPA UB.
- Imasdiani, Purnamasari, I., & Amijaya, F. D. T. (2022). Perbandingan Hasil Analisis Cluster dengan Menggunakan Metode Average Linkage dan Metode Ward (Studi Kasus : Kemiskinan Di Provinsi Kalimantan Timur Tahun 2018). Jurnal Eksponensial, 13(1), 9–18.
- Johnson, R. A. (2009). Applied Multivariate Statistical Analysis. In Pearson International Edition (Sixth Edit, pp. 690–694).
- Laeli, S. (2014). Analisis Cluster dengan Average Linkage Method dan Ward ’s Method untuk Data Responden Nasabah Asuransi Jiwa Unit Link. FMIPA UNY.
- Nafisah, Q., & Chandra, N. E. (2017). Analisis Cluster Average Linkage Berdasarkan Faktor-Faktor Kemiskinan di Provinsi Jawa Timur. Zeta - Math Journal, 3(2), 31–36. <https://doi.org/10.31102/zeta.2017.3.2.31-36>
- Narimawati, U. (2008). Teknik-Teknik Analisis Multivariat untuk Riset Ekonomi. In Graha Ilmu (pp. 14–15).
- Pratama, R. C., & Nasrudin. (2022). Pengelompokan Kabupaten/Kota Di Provinsi Papua Berdasarkan Indikator Kesejahteraan Rakyat 2020 (Clustering Of Regencies in Papua Based on People’s Welfare Indicators 2020). Seminar Nasional Official Statistics , 2020, 853–862.
- Sugiyono. (2013). Metodologi Penelitian Kuantitatif, Kualitatif, dan R&D. ALFABETA.
- Widodo, E., Mashita, S. N., & Prasetyowati, Y. G. (2020). Perbandingan Metode Average Linkage, Complete Linkage, dan Ward’s pada Pengelompokan Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Indikator Indeks Pembangunan Manusia. Faktor Exacta, 13(2), 81. <https://doi.org/10.30998/faktorexacta.v13i2.6581>.