

Penggerombolan provinsi di Indonesia berdasarkan instrumen akreditasi satuan pendidikan jenjang SMK menggunakan *K-means* dan *average linkage*

(Clustering of provinces in Indonesia based on accreditation instruments for vocational high school educational units using K-means and average linkage)

Andina Fahriya, Febryna Sembiring, Budi Susetyo*

Program Studi Statistika dan Sains Data, IPB University

*korespondensi: budisu@apps.ipb.ac.id

Received: 27-06-2023, accepted: 20-09-2024

Abstract

Improvement and updates need to be done in order to maintain the existence of a school. Accreditation is one of the references to assess the excellence of a school. There are several components used in the accreditation assessment included in the IASP, namely Graduate Quality, Learning Process, Teacher Quality, and School Management. Additionally, to determine which provinces have low, medium, or high IASP scores, clustering is performed on the IASP scores of those provinces. Cluster analysis is a method used to group research objects based on similarities in their characteristics. In this study, clustering was performed using the K-means and average linkage methods on the average IASP scores of vocational high schools (SMK) in 34 provinces in Indonesia. With the Elbow Criterion approach, four clusters were formed for each method. The results of Dunn Index showed that the average linkage method performed better in clustering compared to the K-Means method.

Keywords: IASP, Cluster Analysis, K-Means, Average Linkage

MSC2020: 62H30

1. Pendahuluan

Kebijakan dan program pemerintah dalam bidang pendidikan didasarkan pada fungsi dan tujuan pendidikan Indonesia yang termuat dalam Undang-Undang Dasar Republik Indonesia tahun 1945. Pemerintah Indonesia menjamin bahwa setiap warga negara memiliki hak yang sama untuk mendapatkan pendidikan berkualitas sesuai dengan standar pendidikan nasional yang ditetapkan [1]. Pendidikan Indonesia memiliki tujuan mengembangkan dan meningkatkan kualitas kehidupan manusia, produktivitas masyarakat, dan daya saing di kancah internasional. Tujuan lain pendidikan Indonesia adalah terbentuknya generasi muda yang berkarakter agar dapat ikut terlibat dalam memajukan negara Indonesia [2].

Pendidikan merupakan kunci untuk memperbaiki kehidupan di masa depan. Tingkat satuan pendidikan yang dianggap untuk menyelenggarakan pendidikan kejuruan pada jenjang pendidikan menengah sebagai lanjutan dari SMP, MTs, atau bentuk lain sederajat adalah Sekolah Menengah Kejuruan (SMK). Pada umumnya setiap SMK memiliki hasil evaluasi BAN-S/M menunjukkan akreditasi sekolah tersebut. Badan Akreditasi Nasional Sekolah/Madrasah mulai tahun 2018 telah merancang perubahan sistem akreditasi dari perubahan paradigma berbasis keputusan administratif (*compliance*) menjadi berbasis kinerja (*performance*). Kegiatan akreditasi diharapkan sebagai pendorong dan dapat menciptakan suasana kondusif bagi perkembangan pendidikan serta memberikan arahan dalam melakukan penjaminan mutu sekolah yang berkelanjutan, guna mencapai mutu yang diharapkan [3].

Mengingat persaingan sekolah semakin kompetitif, maka suatu sekolah perlu melakukan perbaikan dan pembaharuan yang berkesinambungan agar tetap dapat mempertahankan eksistensi sekolah tersebut. Acuan yang digunakan untuk melihat keunggulan suatu sekolah dilihat dari akreditasi. Yang menjadi komponen utama dalam penilaian akreditasi adalah Mutu Lulusan, Proses Pembelajaran, Mutu Guru dan Manajemen Sekolah [4]. Menurut Kemendikbud [5], Indonesia menempati urutan ke-40 dari 40 negara pada pemetaan kualitas pendidikan. Hal ini ditunjukkan dari salah satu komponen penilaian akreditasi Mutu Guru yang hanya mencapai 44.5 dari standar nilai 75.

Dari permasalahan di atas diperlukan untuk mengelompokkan sekolah mana saja yang perlu perhatian khusus atau yang sudah cukup baik berdasarkan penilaian akreditasi sekolah. Metode yang dapat dalam menyelesaikan permasalahan tersebut adalah melalui salah satu cabang ilmu statistika yaitu memakai analisis gerombolan atau *clustering*. Analisis gerombol merupakan teknik analisis peubah ganda yang dilakukan dengan tujuan menemukan struktur gerombol dalam satu set data yang ditandai dengan kemiripan masing-masing objek di dalam gerombol yang sama, namun memiliki perbedaan yang sangat besar pada masing-masing objek di gerombol yang berbeda [6]. Analisis gerombol merupakan salah satu metode interdependensi (saling ketergantungan), maka dari itu tidak terdapat perbedaan antara peubah respon dan peubah prediktor [7].

Tujuan dari analisis gerombol adalah (1) menyelidiki struktur dasar dari suatu data, (2) klasifikasi, yaitu untuk menentukan tingkat kemiripan antar objek, dan (3) menyimpulkan data masuk ke dalam kelompok-kelompok yang dapat dipahami [8]. Beberapa algoritma yang dapat dilakukan dalam penggerombolan untuk menarik informasi mengenai sifat dari suatu gerombol, seperti algoritma K-Rataan (*K-Means*) dan Pautan Rataan (*Average Linkage*). Data yang memiliki sifat atau ciri yang sama akan dikelompokkan ke dalam gerombol atau *cluster* yang sama dan data yang memiliki sifat atau ciri yang berbeda akan dikelompokkan ke dalam gerombol lain.

Sebelumnya beberapa penelitian mengenai penggerombolan melalui metode *K-Means* dan *Average Linkages*. Penelitian [9] membandingkan metode *K-Means* dan metode

gerombol hirarki pada data segemntasi pelanggan. Berdasarkan nilai *Davis-Bouldin*, *Silhouette* dan *Dunn Index*, diperoleh metode terbaik adalah gerombol *K-Means*. Penelitian [10] membandingkan antara metode gerombol *K-Means* dan *average linkage* pada data kemiskinan di Provinsi Jawa Barat. Metode terbaik yang diperoleh adalah *average linkage* dengan hasil *Silhouette Index* 0.35, sedangkan hasil *Silhouette Index K-means* adalah 0.2. Penelitian lainnya [11] membandingkan penerapan algoritma *K-means* dan *average linkage* pada data industri di Kabupaten Wajo. Diperoleh *K-Means* merupakan metode terbaik dengan nilai centroid yang lebih kecil dibandingkan dengan *average Linkage*.

Pada penelitian ini, akan dilakukan penggerombolan dengan membandingkan metode *K-means* dan *average linkage* pada data Instrumen Akreditasi Satuan Pendidikan (IASP) jenjang Sekolah Menengah Kejuruan (SMK) seluruh provinsi di Indonesia.

2. Metodologi

Penelitian ini menggunakan data rata-rata skor Instrumen Akreditasi Satuan Pendidikan (IASP) jenjang Sekolah Menengah Kejuruan (SMK) pada 34 provinsi di Indonesia. Di dalamnya memuat empat kategori penilaian mutu, yaitu Mutu Lulusan, Proses Pembelajaran, Mutu Guru dan Manajemen Sekolah. Tabel 1 merupakan ringkasan data yang digunakan.

Tabel 1. Rerata skor IASP SMK pada 34 provinsi di Indonesia

No	Provinsi	Rerata Mutu Lulusan	Rerata Proses Pembelajaran	Rerata Mutu Guru	Rerata Manajemen
1	Aceh	73.08	73.15	66.39	70.02
2	Bali	93.75	95.23	91.41	95.27
3	Banten	79.92	80.00	73.33	78.49
4	Bengkulu	83.88	86.11	75.48	84.24
5	DI Yogyakarta	88.36	92.12	85.14	91.38
6	DKI Jakarta	90.15	90.57	89.24	90.69
7	Gorontalo	81.15	86.11	76.00	80.29
⋮	⋮	⋮	⋮	⋮	⋮
34	Sumatera Utara	85.04	85.17	79.49	85.74

2.1 Algoritma K-Rataan

Gerombol K-rataan atau biasa disebut *K-means* merupakan analisis yang digunakan untuk mengelompokkan suatu n objek ke dalam k kelompok. Algoritma penggerombolan *K-means* adalah salah satu teknik penggerombolan non-hirarki yang populer yang digunakan untuk mengidentifikasi kesamaan antara objek berdasarkan vektor jarak. Teknik ini memiliki beberapa keunggulan termasuk kekompakan, efisiensi, dan kecepatan [8]. Berikut merupakan algoritma menggunakan *K-Means* dalam

penggerombolan [12]:

1. Menentukan jumlah gerombol dengan metode *Elbow Criterion* dengan menggunakan Persamaan (1) [13],

$$SSE = \sum_{k=1}^k \sum_{x_i \in S_k} \|N_i - C_k\| \quad (1)$$

2. Tentukan pusat k pada gerombol pertama (*centroid*) yang diambil dari objek secara acak.
3. Menghitung jarak dari setiap objek ke *centroid* menggunakan *Euclidean Distance* [14] seperti pada Persamaan (2).

$$D_{ij} = \sqrt{(x_{1j} - v_{ij})^2 + \dots + (x_{ij} - v_{kj})^2} \quad (2)$$

dengan,

D_{ij} : Jarak objek i ke centroid ke k
 x_{ij} : Objek ke i pada atribut ke j
 v_{kj} : Titik centroid ke k pada atribut ke j

4. Hitung kembali *centroid* dengan objek gerombol yang baru menggunakan Persamaan (3),

$$\mu_k = \frac{1}{N_k} \sum_{i=1}^{N_k} x_i \quad (3)$$

dengan,

μ_k : *Centroid* gerombol ke k
 N_k : Banyaknya objek pada gerombol ke k
 x_i : Objek ke i pada gerombol ke k

5. Jika *centroid* tidak berubah, maka proses penggerombolan sudah selesai. Jika *centroid* masih berubah, maka ulangi langkah 3 sampai diperoleh *centroid* yang sama.

2.2 Algoritma Pautan Rataan

Gerombol pautan rata-rata atau biasa disebut *average linkage* merupakan salah satu metode gerombol hirarki yang menggunakan konsep jarak rata-rata antar objek pada satu gerombol dengan objek lain pada gerombol lain. Berikut merupakan algoritma *average linkage* dalam penggerombolan [15]:

1. Menentukan jumlah gerombol dengan metode *Elbow Criterion* seperti pada Persamaan (1).
2. Menentukan objek bersesuaian yang memiliki jarak terdekat menggunakan matriks jarak pada Persamaan (4) [16]:

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (4)$$

dengan,

p : Jumlah peubah yang dianalisis

x_{ik} : Objek ke i pada atribut ke k

x_{jk} : objek ke j pada atribut ke k

3. Menggabungkan objek tersebut menjadi satu gerombol, misalnya gerombol (AB).
4. Menghitung jarak antar gerombol (AB) dengan objek lain yang belum tergabung, misalnya C menggunakan Persamaan (5) [17]

$$d_{(AB)C} = \frac{\sum_u \sum_v d_{uv}}{N_{(AB)} N_C} \quad (5)$$

dengan,

d_{uv} : Jarak antar objek u pada gerombol AB dan objek v pada gerombol C

$N_{(AB)}$: Jumlah objek pada gerombol AB

N_C : Jumlah objek pada gerombol C

dalam hal ini, semakin kecil jarak antar objek, maka tingkat kemiripan objek tersebut semakin tinggi, begitupun sebaliknya.

5. Mengulang langkah 2 sampai tergabung menjadi satu gerombol.

2.3 Indeks Dunn

Pengujian validitas dilakukan dengan membandingkan metode *K-means* dan *average linkage* untuk memperoleh metode yang lebih optimum. Pengujian dilakukan dengan menggunakan nilai Indeks *Dunn*, di mana nilai ini memiliki interval antara $0 - \infty$. Indeks *Dunn* dapat dihitung dengan mengukur jarak terbesar antar gerombol dengan terkecil di dalam gerombol [18]. Jika nilai *Dunn Index* semakin besar, maka hasil penggerombolan semakin baik. Nilai *Dunn Index* dapat dihitung menggunakan Persamaan (6) [15],

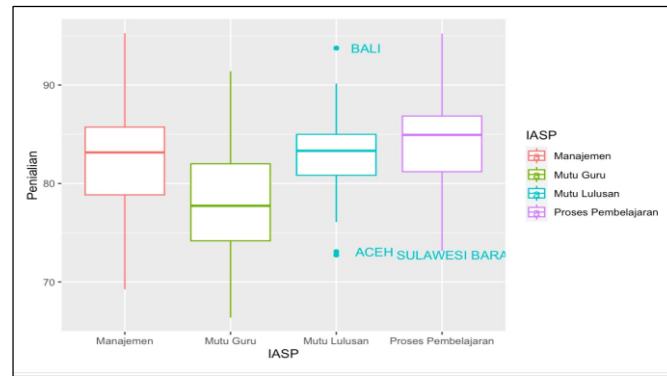
$$D(C) = \frac{\min_{C_k, C_l \in C, C_k \neq C_l} \left(\min_{i \in C_k, j \in C_l} d_{ij} \right)}{\max_{C_m \in C} \text{diam}(C_m)}. \quad (6)$$

3. Hasil dan Pembahasan

3.1 Eksplorasi Data

Eksplorasi data dilakukan untuk melihat karakteristik data yang digunakan. Gambar 1 merupakan eksplorasi data menggunakan rata-rata IASP SMK 34 provinsi di Indonesia.

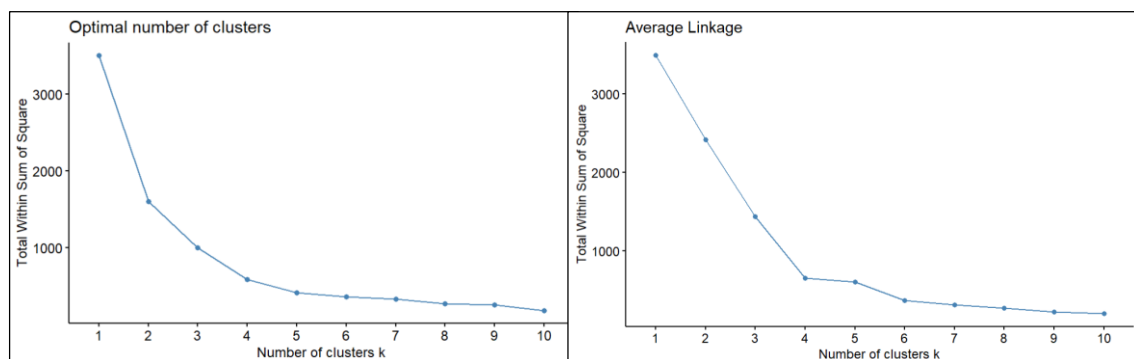
Berdasarkan *Boxplot* di atas, terdapat pencilan pada indikator Mutu Lulusan, yaitu Bali, Aceh dan Nusa Tenggara barat. Pencilan atas, yaitu provinsi Bali menandakan nilai Mutu Lulusan yang sangat tinggi dibandingkan daerah lainnya. Sedangkan pencilan bawah menandakan kurangnya Mutu Lulusan pada provinsi Aceh dan Nusa tenggara barat dibandingkan daerah lainnya.



Gambar 1. Sebaran IASP 34 Provinsi Indonesia

3.2 Analisis Gerombol

Dilakukan analisis pada data hasil penilaian akreditasi jenjang Sekolah Menengah Kejuruan (SMK) seluruh provinsi di Indonesia. Sebelum dilakukan penggerombolan pada data, perlu dicari terlebih dahulu jumlah gerombol yang akan digunakan pada gerombol *K-Means* dan *average linkage* sebagai berikut.



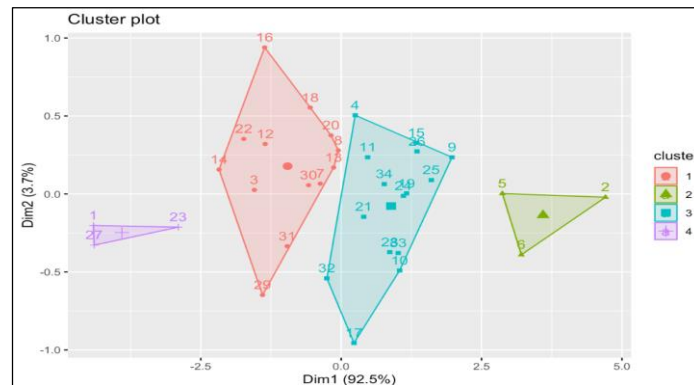
Gambar 2. Jumlah gerombol

Berdasarkan Gambar 2 baik metode *K-means* maupun *average linkage* mengalami pelandaian pada gerombol ke-5, sehingga diperoleh jumlah optimum gerombol *K-means* dan *average linkage* adalah sebanyak 4 gerombol.

3.3 Gerombol K-Rataan

Gerombolan K-Rataan (*K-Means*) dengan jumlah gerombol sebanyak 4 dilakukan menggunakan *software Rstudio*. Gambar 3 adalah plot gerombol *K-Means* yang diperoleh.

Berdasarkan Gambar 3, gerombol 1 mempunyai 13 provinsi, gerombol 2 mempunyai 3 provinsi, gerombol 3 mempunyai 15 provinsi dan gerombol 4 mempunyai 3 provinsi. Data masing-masing gerombol yang telah memuat provinsi di dalamnya dapat dilihat pada Tabel 2.



Gambar 3. Hasil Plot Gerombol *K-Means*

Tabel 2. Hasil analisis gerombol *K-means*

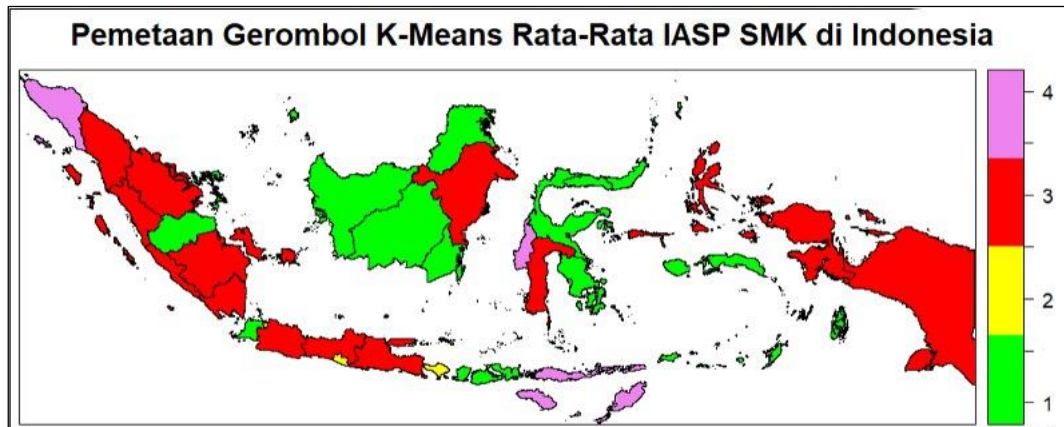
Provinsi	Gerombol
Banten, Gorontalo, Jambi, Kalimantan Barat, Kalimantan Selatan, Kalimantan Tengah, Kalimantan Utara, Kepulauan Riau, Maluku, Nusa Tenggara Barat, Sulawesi Tengah, Sulawesi Tenggara, Sulawesi Utara	1
Bali, DI Yogyakarta, DKI Jakarta	2
Bengkulu, Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Timur, Kepulauan Bangka Belitung, Lampung, Maluku Utara, Papua, Papua Barat, Riau, Sulawesi Selatan, Sumatera Barat, Sumatera Selatan, Sumatera Utara	3
Aceh, Nusa Tenggara Timur, Sulawesi Barat	4

Selanjutnya dilakukan identifikasi karakteristik gerombol untuk memberikan gambaran terkait penggerombolan di atas, sebagai berikut:

Tabel 3. Rata-Rata Skor IASP dengan gerombol *K-means*

Gerombol	Rata-Rata Mutu Lulusan	Rata-Rata Proses Pembelajaran	Rata-Rata Mutu Guru	Rata-Rata Manajemen
1	81.65	82.07	74.51	79.32
2	90.75	92.64	88.59	92.44
3	84.61	86.54	80.49	85.08
4	73.97	74.44	67.95	71.53

Pada gerombol 1 sampai 4, terlihat bahwa Mutu Guru merupakan indikator dengan nilai rata-rata terendah, sementara Proses Pembelajaran merupakan indikator dengan nilai rata-rata tertinggi. Berikut merupakan gambar penggerombolan *K-Means* menggunakan visualisasi map.

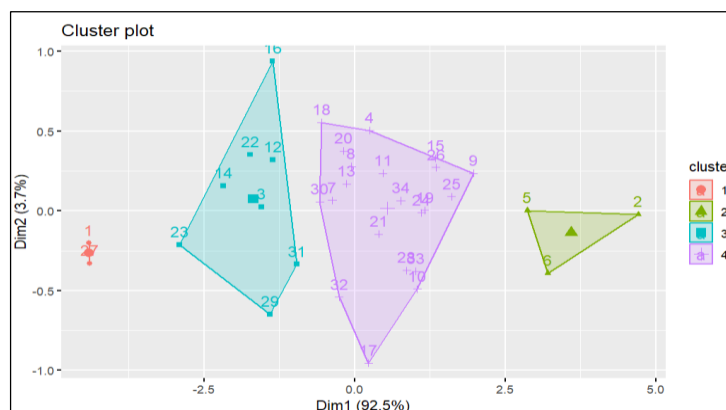


Gambar 4. Peta Indonesia berdasarkan gerombol *K-means*

Jika dilihat dari nilai rata-rata IASP provinsi tiap gerombol, gerombol 1 merupakan provinsi dengan penilaian IASP yang cukup tinggi. Gerombol 2 merupakan provinsi dengan rata-rata IASP tertinggi diantara gerombol lainnya, artinya upaya dalam meningkatkan akreditasi sekolah sudah gencar dilakukan. Gerombol 3 merupakan kelompok provinsi yang juga sudah termasuk baik karena penilaian IASP hampir sama dengan rata-rata penilaian IASP untuk jenjang SMK di Indonesia, sedangkan gerombol 4 merupakan provinsi dengan nilai rata-rata IASP yang paling rendah dibanding kelompok provinsi yang lain, artinya upaya dalam meningkatkan akreditasi sekolah belum baik dan perlu perhatian khusus untuk provinsi di dalamnya.

3.4 Gerombol Pautan Rataan

Penggerombolan dengan Pautan Rataan atau *average linkage* dilakukan menggunakan bantuan *software Rstudio*. Sebelumnya telah diperoleh jumlah gerombol optimum menghasilkan 4 gerombol. Berikut merupakan plot gerombol *average linkage* yang diperoleh:



Gambar 5. Hasil plot gerombol *average linkage*

Gerombol 1 mempunyai 2 provinsi, gerombol 2 mempunyai 3 provinsi, gerombol 3 mempunyai 8 provinsi dan gerombol 4 mempunyai 20 provinsi. Data masing-masing gerombol yang telah memuat provinsi di dalamnya dapat dilihat pada Tabel 3. berikut:

Tabel 4. Hasil analisis gerombol *average linkage*

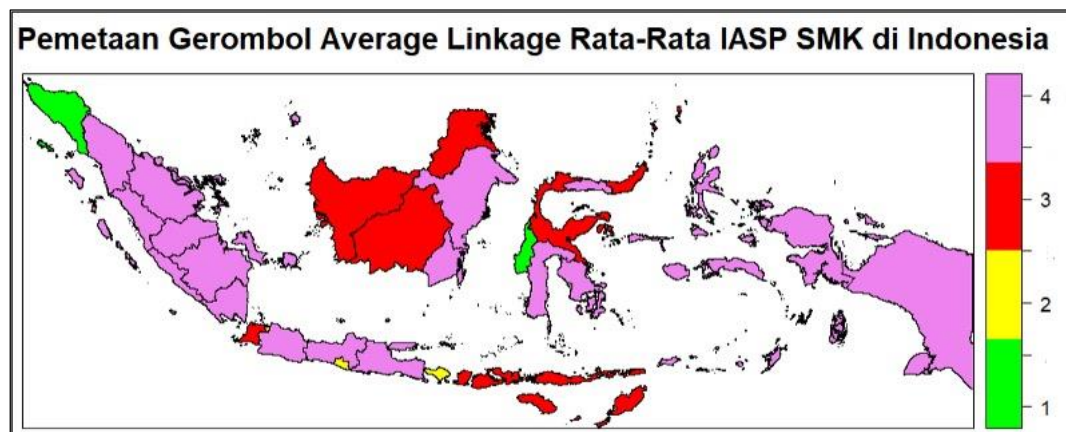
Provinsi	Gerombol
Aceh, Sulawesi Barat	1
Bali, Di Yogyakarta, DKI Jakarta	2
Banten, Kalimantan Barat, Kalimantan Tengah, Kalimantan Utara, Nusa Tenggara Barat, Nusa Tenggara Timur, Sulawesi Tengah, Sulawesi Utara	3
Bengkulu, Gorontalo, Jambi, Jawa Barat, Jawa Tengah, Jawa Timur, Kalimantan Selatan, Kalimantan Timur, Kepulauan Bangka Belitung, Kepulauan Riau, Lampung, Maluku, Maluku Utara, Papua, Papua Barat, Riau, Sulawesi Selatan, Sulawesi Tenggara, Sumatera Barat, Sumatera Selatan, Sumatera Utara	4

Selanjutnya, dilakukan identifikasi karakteristik gerombol untuk memberikan gambaran terkait penggerombolan di atas sebagai berikut:

Tabel 5. Rata-rata skor IASP dengan gerombol *average linkage*

Gerombol	Rata-Rata Mutu Lulusan	Rata-Rata Proses Pembelajaran	Rata-Rata Mutu Guru	Rata-Rata Manajemen
1	72.91	73.41	66.76	69.64
2	90.75	92.64	88.59	92.44
3	80.16	79.56	73.08	77.45
4	84.07	85.95	79.12	83.96

Pada gerombol 1, 2 dan 4, terlihat bahwa Mutu Guru merupakan indikator dengan nilai rata-rata terendah, sementara Proses Pembelajaran merupakan indikator dengan nilai rata-rata tertinggi. Kemudian untuk gerombol 3, Mutu Guru merupakan indikator dengan nilai rata-rata terendah, sedangkan Mutu Lulusan merupakan indikator dengan rata-rata tertinggi. Berikut merupakan gambar penggerombolan *average linkage* menggunakan visualisasi map.



Gambar 6. Peta Indonesia berdasarkan gerombol *average linkage*

Jika dilihat dari nilai rata-rata IASP provinsi tiap gerombol, gerombol 1 merupakan provinsi dengan penilaian IASP terendah dibandingkan kelompok provinsi yang lain. Hal

ini menunjukkan bahwa perlu perhatian khusus untuk provinsi yang ada pada gerombol tersebut dan upaya dalam meningkatkan akreditasi belum cukup gencar dilakukan. Gerombol 2 merupakan provinsi dengan rata-rata IASP tertinggi diantara gerombol lainnya, artinya bahwa kelompok provinsi tersebut sudah cukup baik dalam meningkatkan akreditasi sekolah. Gerombol 3 merupakan kelompok provinsi dengan rata-rata IASP masih dibawah rata-rata dari IASP Indonesia dan masih lebih baik dibandingkan dengan gerombol 1, artinya upaya dalam meningkatkan akreditasi sudah cukup baik. Sedangkan gerombol 4 merupakan provinsi dengan nilai rata-rata IASP yang kurang baik karena indikator mutu guru pada gerombol tersebut masih dibawah rata-rata. Ini menunjukkan bahwa perlu perhatian lebih lagi dalam peningkatan akreditasi sekolah pada kelompok tersebut.

3.5 Indeks Dunn

Setelah dilakukan penggerombolan dengan metode *K-Means* dan *average linkage* dengan jumlah gerombol sebanyak 4, selanjutnya dilakukan perbandingan nilai menggunakan *Dunn Index* untuk mengetahui gerombol mana yang lebih optimum. Berikut nilai *Dunn Index* yang diperoleh:

Tabel 6. Nilai *dunn index*

Metode Gerombol	<i>Dunn Index</i>
Metode <i>Average Linkage</i>	0.31
Metode <i>K-Means</i>	0.19

Berdasarkan Tabel 6, nilai *Dunn Index* pada gerombol *average linkage* menghasilkan nilai yang lebih besar dibandingkan dengan gerombol *K-Means*. Hal ini menandakan bahwa metode *average linkage* lebih baik dalam penggerombolan data rata-rata skor IASP SMK pada 34 provinsi di Indonesia.

4. Kesimpulan

Berdasarkan *Elbow Criterion*, metode *K-Means* dan *average linkage* menghasilkan jumlah penggerombolan optimum sebanyak 4 gerombol. Untuk metode *K-means*, gerombol 1 mempunyai 13 provinsi, gerombol 2 mempunyai 3 provinsi, gerombol 3 mempunyai 15 provinsi dan gerombol 4 mempunyai 3 provinsi. Untuk metode *Average Linkage*, gerombol 1 mempunyai 2 provinsi, gerombol 2 mempunyai 3 provinsi, gerombol 3 mempunyai 8 provinsi dan gerombol 4 mempunyai 20 provinsi. Hasil gerombol untuk kedua metode menunjukkan bahwa Mutu Guru merupakan indikator dengan nilai rata-rata terendah, sehingga perlu adanya peningkatan pada indikator tersebut. Berdasarkan nilai *Dunn Index*, metode *average linkage* merupakan metode terbaik dalam menggerombolkan rata-rata IASP Provinsi Indonesia. Karakteristik gerombol yang terbentuk menggunakan metode *average linkage* yaitu gerombol 1 termasuk provinsi dengan rata-rata IASP terendah, hal ini bisa menjadi masukan pada SMK di provinsi Aceh dan Sulawesi Barat untuk dapat meningkatkan kualitas sekolah,

terutama dalam hal indikator penilaian akreditasi sekolah. Kemudian gerombol 2 merupakan provinsi dengan rata-rata IASP tertinggi, hal ini menandakan kualitas SMK di tiga provinsi tersebut yang unggul dan memiliki kualitas khususnya dalam indikator penilaian akreditasi sekolah.

Daftar Pustaka

- [1] B. Susetyo, S. P. Soetantyo, M. Sayuti, and D. Nur, "The innovation and the transformation of indonesian schools accreditation management system," *Indones. J. Learn. Adv. Educ.*, vol. 4, no. 2, pp. 128–139, 2022. [\[CrossRef\]](#)
- [2] R. Lase, B. Susetyo, and U. Syafitri, "Application of multivariate exploration analysis of the vocational high school accreditation data in indonesia," *Int. J. Sci. Eng. Res.*, vol. 7, no. 12, 2016. [\[CrossRef\]](#)
- [3] Badan Akreditasi Nasional, *Pedoman Akreditasi Sekolah dan Madrasah Tahun 2023*. Jakarta: Badan Akreditasi Nasional Sekolah/Madrasah, 2023.
- [4] Badan Akreditasi Nasional Sekolah/Madrasah, *Instrumen Akreditasi Satuan Pendidikan 2020 Jenjang Sekolah Menengah Atas/Madrasah Aliyah*. Jakarta: Badan Akreditasi Nasional Sekolah/Madrasah, 2020.
- [5] S.S. Handoyo, M. Yusro, and A.A. Jumhur, "Akreditasi SMK/MAK Sebagai Bentuk Akuntabilitas Publik dan Penjaminan Mutu Pendidikan Kejuruan," 2016. [\[CrossRef\]](#)
- [6] K.P. Sinaga and M.-S. Yang, "Unsupervised k-means clustering algorithm," *IEEE Xplore*, vol. 8, pp. 80716–80727, 2020. [\[CrossRef\]](#)
- [7] A.A. Mattjik and I.M. Sumertajaya, *Sidik Peubah Ganda Dengan Menggunakan SAS*. Bogor: IPB PRESS, 2011.
- [8] D. Abdullah, S. Susilo, A.S. Ahmar, R. Rusli, and R. Hidayat, "The application of k-means clustering for province clustering in indonesia of the risk of the covid-19 pandemic based on covid-19 data," *Qual. Quant.*, vol. 56, pp. 1283–1291, 2021. [\[CrossRef\]](#)
- [9] A. Abdulhafedh, "Incorporating k-means, hierarchical clustering and pca in customer segmentation," *J. City Dev.*, vol. 3, no. 1, pp. 12–30, 2021. [\[CrossRef\]](#)
- [10] D. Widyadhana, R.B. Hastuti, I. Kharisudin, and F. Fauzi, "Perbandingan analisis klaster k-means dan average linkage untuk pengklasteran kemiskinan di provinsi jawa tengah," *Prism. Pros. Semin. Nas. Mat.*, vol. 4, pp. 584–594, 2021. [\[CrossRef\]](#)

- [11] A. Zikir, K. Nurfadilah, Irwan, and Adiatma, “Perbandingan metode clustering dengan menggunakan metode average linkage dan metode k-means pada industri kecil dan menengah di kabupaten wajo,” *J. Mat. dan Stat. serta Apl.*, vol. 10, no. 2, pp. 57–62, 2022. [\[CrossRef\]](#)
- [12] N. Butarbutar, A.P. Windarto, D. Hartama, and Solikhun, “Komparasi kinerja algoritma fuzzy c-means dan k-means dalam pengelompokan data siswa berdasarkan prestasi nilai akademik siswa,” *JURASIK (Jurnal Ris. Sist. Inf. Tek. Inform.*, vol. 1, no. 1, pp. 46–55, 2016, doi: 10.30645/jurasik.v1i1.8. [\[CrossRef\]](#)
- [13] N.T. Hartanti, “Metode elbow dan k-means guna mengukur kesiapan siswa smk dalam ujian nasional,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 6, no. 2, pp. 82–89, 2020. [\[CrossRef\]](#)
- [14] Nurahman, A. Purwanto, and S. Mulyanto, “Klasterisasi sekolah menggunakan algoritma k-means berdasarkan fasilitas, pendidik, dan tenaga pendidik,” *Matrik J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 2, pp. 337–349, 2022. [\[CrossRef\]](#)
- [15] A. Wijaya, Nurhasanah, F. AR, and A. Rusyana, “Perbandingan metode gerombol pautan lengkap dan pautan rata-rata untuk pengelompokan kemiskinan kabupaten/kota di indonesia,” *J. Data Anal.*, vol. 3, no. 1, pp. 13–25, 2020. [\[CrossRef\]](#)
- [16] G.L. Koesnadi *et al.*, “Pengelompokan provinsi berdasarkan prioritas potensi sektor maritim indikator blue economy menggunakan analisis cluster average linkage,” *J. Sains Mat. dan Stat.*, vol. 9, no. 1, pp. 65–77, 2023. [\[CrossRef\]](#)
- [17] R.A. Johnson and D.W. Wichren, *Applied Multivariate Statistical Analysis 6th Edition*. New Jersey: Pearson Education, 2007.
- [18] A.D. Savitri, F.A. Bachtiar, and N.Y. Setiawan, “Segmentasi pelanggan menggunakan metode k-means clustering berdasarkan model rfm pada klinik kecantikan (studi kasus : belle crown malang),” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 9, pp. 2957–2966, 2018. [\[CrossRef\]](#)