

Klasterisasi profil sekolah penerima bantuan revitalisasi SMK 2025 berbasis K-Medoids dan *Gower Distance*

Arie Wibowo Khurniawan¹, Gustriza Erda², Tengku Irfan Wira Buana², Tengku Alfaisal², Arisman Adnan², Susilawati³, Miftahus Sholih⁴

¹ Direktorat Sekolah Menengah Kejuruan, Kementerian Pendidikan Dasar dan Menengah, Jakarta, 10270, Indonesia

² Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Riau, Riau, 28293, Indonesia

³ Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Riau, Riau, 28293, Indonesia

⁴ Program Studi Statistika, Universitas Sultan Ageng Tirtayasa, Banten, 42163, Indonesia

Korespondensi

Gustriza Erda

gustrizaerda@lecturer.unri.ac.id

Program Studi Statistika,
Fakultas Matematika dan Ilmu
Pengetahuan Alam, Universitas
Riau, Riau, 28293, Indonesia

Artikel Histori

Submitted: 20 Mei 2026

Revised Paper: 28 Mei 2026

Accepted: 28 Mei 2026

Published : 30 Mei 2026

© 2026, Authors



This work is licensed under a
Creative Commons
Attribution-ShareAlike 4.0
International License

Abstrak

Purpose. Penelitian ini bertujuan memetakan profil sekolah penerima bantuan berdasarkan 12 variabel multidimensi serta merumuskan rekomendasi sebagai dasar kebijakan revitalisasi berikutnya. **Method.** Penelitian ini menggunakan metode klasterisasi dengan *Gower Distance* untuk menangani data campuran kategorik dan numerik, serta algoritma K-Medoids dengan implementasi *FasterPAM*. Jumlah kluster optimal ditentukan melalui kombinasi *Hierarchical Clustering* dengan *Ward Linkage* dan *Silhouette Score* yang kemudian divisualisasikan dengan *Multidimensional Scaling*. **Finding.** Hasil analisis mengidentifikasi bahwa terdapat tiga kluster sekolah penerima bantuan, yaitu Kluster 1 berisi 454 sekolah (31%) sebagai kelompok berpengalaman yang terkonsentrasi di Pulau Jawa, Kluster 2 berisi 600 sekolah (41%) dengan alokasi anggaran tertinggi yang menjangkau wilayah 3T dan menunjukkan proporsi pelaksanaan tahun tunggal yang relatif baik, serta Kluster 3 berisi 411 sekolah (28%) sebagai penerima pertama kali yang teridentifikasi sebagai kelompok prioritas untuk pendampingan intensif. **Practical Implication.** Temuan penelitian menunjukkan bahwa selain besaran anggaran, faktor pengalaman institusional sekolah dan distribusi penjadwalan operasional turut berperan penting dalam mendukung kelancaran pelaksanaan program. **Originality.** Hasil penelitian ini memberikan dasar empiris bagi pengembangan kebijakan revitalisasi SMK yang lebih terdiferensiasi dan adaptif ke depan.

Kata Kunci : *Gower Distance*; Kluster; K-Medoids *Fasterpam*; Revitalisasi SMK.

Abstract

Purpose. This study aims to map the profile of aid recipient schools based on 12 multidimensional variables and formulate recommendations as a basis for subsequent revitalization policies. **Method.** This study uses the clustering method with *Gower Distance* to handle mixed categorical and numeric data, as well as the K-Medoids algorithm with *FasterPAM* implementation. The optimal number of clusters is determined through a combination of *Hierarchical Clustering* with *Ward Linkage* and *Silhouette Score* which is then visualized with *Multidimensional Scaling*. **Findings.** The results of the analysis identified that there are three clusters of aid recipient schools, namely Cluster 1 containing 454 schools (31%) as an experienced group concentrated in Java Island, Cluster 2 containing 600 schools (41%) with the highest budget allocation reaching the 3T region and showing a relatively good proportion of single-year implementation, and Cluster 3 containing 411 schools (28%) as first-time recipients identified as a priority group for intensive mentoring. **Practical Implications.** The research findings show that in addition to the budget size, the institutional experience of the school and the distribution of operational scheduling also play an important role in supporting the smooth implementation of the program. **Originality.** The results of this study provide an empirical basis for the development of a more differentiated and adaptive vocational high school revitalization policy in the future.

Keywords: Cluster Analysis; *Gower Distance*; K-Medoids *Fasterpam*; Revitalization of Vocational High Schools.

How to cite:

Khurniawan, A. W., Erda, G., Wira, B. T. I., Tengku, A., Arisman, A., Susilawati, & Miftahus, S. (2026). Klasterisasi profil sekolah penerima bantuan revitalisasi SMK 2025 berbasis K-Medoids dan *Gower Distance*. *Vocational Education Policy*, 1(1), 1–19. <https://doi.org/10.66599/wp.v1i1.40>

PENDAHULUAN

Sekolah Menengah Kejuruan (SMK) merupakan salah satu lembaga pendidikan yang memiliki peran penting dalam menyiapkan lulusan yang siap memasuki dunia kerja. SMK dirancang untuk membekali peserta didik dengan keterampilan teknis sesuai bidang keahlian masing-masing, sehingga lulusan diharapkan mampu berkompetisi, beradaptasi, serta mengikuti perkembangan teknologi dan kebutuhan industri yang terus berubah. Namun pada pelaksanaannya, SMK mengalami tantangan, khususnya dalam hal pemenuhan dan pemerataan sarana dan prasarana pendidikan yang layak. Data Pokok Pendidikan (Dapodik) per April 2025 mencatat bahwa dari total 14.490 SMK di Indonesia, terdapat ruang kelas yang mengalami kerusakan dengan rincian 4.825 ruang kelas mengalami kerusakan berat, 6.835 ruang kelas mengalami kerusakan sedang, dan 9.908 ruang kelas mengalami kerusakan ringan. Angka-angka ini bukan sekadar statistik infrastruktur, melainkan cerminan tantangan dalam pemenuhan hak konstitusional warga negara atas pendidikan bermutu sebagaimana diamanatkan Undang-Undang Dasar 1945.

Sebagai respons, Pemerintah menetapkan revitalisasi sebagai salah satu Program Hasil Terbaik Cepat (PHTC) Presiden untuk mendukung Asta Cita ke-4 dan Transformasi Birokrasi dalam RPJMN 2025–2029 dengan proyeksi kebutuhan pendanaan Rp571,6 triliun selama lima tahun. Revitalisasi satuan pendidikan ini diatur di Instruksi Presiden Nomor 9 Tahun 2016 tentang Revitalisasi SMK dan pada Peraturan Presiden Nomor 68 Tahun 2022 tentang Revitalisasi Pendidikan Vokasi dan Pelatihan Vokasi. Pada level operasional, Direktorat SMK melalui Subdirektorat Fasilitasi Sarana Prasarana dan Tata Kelola menyalurkan bantuan kepada 1.465 SMK di 38 provinsi pada Tahun Anggaran 2025 dengan total pagu Rp2,6 triliun. Program ini memiliki karakteristik unik berupa keterlibatan dua skema pendampingan (Perguruan Tinggi dan Zeni TNI), jangkauan ke wilayah Terdepan, Terluar, dan Tertinggal (3T), serta pemberian bantuan kepada sekolah baru maupun penerima berulang.

Skala program yang besar dan keragaman penerima bantuan memunculkan tantangan tersendiri dalam pengelolaan yang efektif. Profil SMK penerima menampilkan heterogenitas multi-dimensi yang signifikan, mulai dari dimensi geografis (Jawa hingga Papua, perkotaan hingga 3T), institusional (negeri atau swasta), historis (penerima baru atau berulang), finansial (anggaran Rp233 juta hingga Rp8,24 miliar), hingga operasional (urutan bimbingan teknis ke-1 hingga ke-11). Heterogenitas ini mengisyaratkan bahwa pendekatan *one-size-fits-all* berpotensi tidak optimal karena setiap tipe sekolah membutuhkan strategi pendampingan yang berbeda. Sistem pemantauan dan evaluasi yang sudah berjalan dengan indikator agregat nasional dapat diperkaya dengan pendekatan multi-dimensi yang menangkap kompleksitas profil penerima secara lebih komprehensif. Pengembangan pemetaan profil penerima bantuan berbasis

karakteristik multi-dimensional dapat menjadi kontribusi metodologis yang melengkapi sistem monitoring eksisting dan mendukung perumusan kebijakan berbasis bukti.

Dalam menjawab tantangan tersebut, penelitian ini menggunakan pendekatan klusterisasi sebagai kerangka analitis. Klusterisasi merupakan teknik *unsupervised learning* yang memungkinkan pemetaan profil entitas berdasarkan kemiripan karakteristik multi-dimensi tanpa label kelas yang ditentukan di awal (Han et al., 2022). Namun, data penelitian berupa data campuran yang terdiri atas variabel nominal, biner, ordinal, dan numerik kontinu membuat metode konvensional seperti K-Means berbasis *Euclidean Distance* tidak sesuai. *Gower Distance* (Gower, 1971) dipilih sebagai solusi karena mampu menangani data campuran melalui perhitungan jarak parsial sesuai tipe variabel (Liu et al., 2024). Adapun K-Medoids dengan implementasi *FasterPAM* (Schubert & Rousseeuw, 2021) dipilih sebagai algoritma klusterisasi karena lebih robust terhadap pencilon dibandingkan K-Means, suatu keunggulan yang relevan mengingat 7,51% data anggaran teridentifikasi sebagai pencilon pada uji IQR.

Beberapa penelitian terdahulu di Indonesia telah menerapkan klusterisasi untuk konteks pendidikan, seperti (Erda et al., 2023) untuk klusterisasi Indeks Pembangunan Manusia provinsi, dan studi terbaru pada jurnal (Vidyananta & Dermawan, 2025) untuk klusterisasi sekolah di Provinsi Riau. Namun, penelitian-penelitian tersebut umumnya menggunakan variabel numerik dan belum menerapkan *Gower Distance* untuk data campuran pada konteks evaluasi program bantuan pendidikan vokasi berskala nasional. Penelitian ini menawarkan kebaruan berupa kombinasi metodologis *Gower Distance* dan K-Medoids *FasterPAM* pada cakupan nasional variabel multi-dimensi sebagai suatu pendekatan yang belum banyak diterapkan dalam evaluasi program bantuan pendidikan vokasi di Indonesia.

Berdasarkan uraian tersebut, penelitian ini dimaksudkan untuk menjawab tiga pertanyaan utama. Pertama, bagaimana profil kluster sekolah penerima bantuan revitalisasi SMK 2025. Kedua, karakteristik khas apa yang membedakan masing-masing kluster ditinjau dari distribusi geografis, anggaran, pendampingan, riwayat bantuan, dan durasi pelaksanaan program. Ketiga, rekomendasi kebijakan terdiferensiasi seperti apa yang dapat dirumuskan berdasarkan profil tiap kluster. Harapannya penelitian ini dapat memberikan dasar empiris bagi pengembangan kebijakan revitalisasi SMK yang lebih tepat sasaran, adaptif terhadap karakteristik penerima bantuan, serta mampu meningkatkan efektivitas pelaksanaan program pada periode berikutnya.

METODE

RANCANGAN PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan desain deskriptif-eksploratif berbasis analisis kluster. Pendekatan ini dipilih karena memungkinkan segmentasi unit observasi berdasarkan

kemiripan multivariat tanpa label kategori yang ditentukan sebelumnya, sehingga relevan untuk merumuskan tipologi sekolah penerima bantuan sebagai dasar perumusan kebijakan berbasis bukti.

SUMBER DAN KARAKTERISTIK DATA

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari database Subdirektorat Sarana, Prasarana, dan Tata Kelola, Direktorat Jenderal Pendidikan Vokasi, Kementerian Pendidikan Dasar dan Menengah (Kemendikdasmen) tahun 2025. Dataset terdiri atas 1.465 SMK penerima bantuan revitalisasi SMK tahun anggaran 2025 yang tersebar di 38 provinsi di Indonesia dan 1 Sekolah Indonesia di Luar Negeri (SILN) yang berlokasi di Malaysia. Setiap baris data memiliki 12 variabel yang merepresentasikan karakteristik geografis, institusional, dan pelaksanaan program bantuan.

LANGKAH Pengerjaan

Eksplorasi data dilakukan untuk memahami struktur, distribusi, dan kualitas data sebelum proses analisis, mencakup perhitungan statistik deskriptif untuk variabel numerik (rerata, median, standar deviasi, minimum, maksimum) dan distribusi frekuensi untuk variabel kategorik. Variabel numerik (Nilai Anggaran) kemudian dinormalisasi menggunakan Min-Max Scaling ke rentang $[0, 1]$ untuk menyeimbangkan kontribusinya terhadap matriks jarak dibandingkan variabel kategorik, dengan formula:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Setelah normalisasi, data fitur akhir berukuran 1.465×12 dengan seluruh variabel ternormalisasi. Matriks jarak *Gower* kemudian dibentuk untuk mengukur kemiripan antar sekolah berdasarkan 12 variabel multi-dimensi. *Gower Distance* (Gower, 1971) dipilih karena mampu menangani data campuran dengan menghitung jarak parsial secara berbeda sesuai tipe variabel kemudian menggabungkannya melalui pembobotan rata-rata. Pendekatan ini lebih sesuai dibanding jarak *Euclidean* yang mengasumsikan seluruh variabel bertipe numerik, sekaligus menghindari pembengkakan dimensi akibat *one-hot encoding* pada variabel berkardinalitas tinggi. Formula umum *Gower Distance* antara dua observasi i dan j dirumuskan sebagai:

$$d_{ij} = \frac{\sum_{k=1}^p w_{ijk} \cdot d_{ijk}}{\sum_{k=1}^p w_{ijk}}$$

dengan p adalah jumlah variabel, w_{ijk} adalah bobot variabel ke- k untuk pasangan (i, j) , dan d_{ijk} adalah jarak parsial pada variabel ke- k . Perhitungan jarak parsial dilakukan dengan ketentuan berikut. Untuk variabel kategorik, digunakan *simple matching coefficient* yang bernilai 0 jika kedua observasi memiliki kategori sama dan 1 jika berbeda. Untuk variabel numerik, digunakan selisih absolut yang dinormalisasi terhadap range variabel tersebut, dirumuskan sebagai:

$$d_{ijk} = \frac{|x_{ik} - x_{jk}|}{R_k}$$

dengan R_k adalah range variabel ke-k. Hasil perhitungan menghasilkan matriks jarak simetris dengan nilai jarak berada pada rentang $[0, 1]$.

Penentuan jumlah kluster optimal dilakukan melalui kombinasi dua pendekatan. Pendekatan visual menggunakan *Hierarchical Clustering* dengan *Ward Linkage* yang diterapkan terhadap matriks jarak *Gower* untuk membangun dendrogram. Metode *Ward* dipilih karena meminimalkan variansi total di dalam kluster sehingga menghasilkan kluster yang kompak dan berukuran seimbang. Dari visualisasi dendrogram, dilakukan inspeksi visual terhadap lompatan jarak antar level untuk mengidentifikasi titik potong yang menghasilkan struktur hierarki yang jelas. Pendekatan kuantitatif menggunakan *Silhouette Score* untuk rentang $K = 2$ hingga $K = 6$, dirumuskan sebagai (Rousseeuw, 1987):

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

dengan $a(i)$ adalah rata-rata jarak observasi i terhadap seluruh objek dalam klasternya sendiri, dan $b(i)$ adalah rata-rata jarak observasi i terhadap seluruh objek pada kluster terdekat. *Silhouette Score* berada pada rentang $[-1, 1]$ dengan nilai mendekati 1 mengindikasikan pengelompokan yang baik. Pemilihan K optimal dilakukan dengan mempertimbangkan kombinasi nilai *Silhouette Score* dan interpretabilitas hasil dalam konteks kebijakan.

Finalisasi kluster dilakukan menggunakan algoritma K-Medoids dengan implementasi *FasterPAM* (*Faster Partitioning Around Medoids*) yang dikembangkan oleh (Schubert & Rousseeuw, 2021). K-Medoids dipilih dibandingkan K-Means karena K-Medoids menggunakan titik data asli sebagai pusat kluster (medoid) sehingga lebih *robust* terhadap pencilon. Selain itu, K-Medoids dapat bekerja langsung dengan matriks jarak *Gower*, sedangkan K-Means membutuhkan koordinat numerik dalam ruang *Euclidean*. Algoritma *FasterPAM* bekerja melalui empat langkah yang diulang hingga konvergen, yaitu inisialisasi K medoid awal menggunakan metode BUILD, alokasi observasi ke medoid terdekat berdasarkan matriks jarak *Gower*, pertukaran medoid jika menurunkan total biaya intra-kluster, dan iterasi hingga tidak ada pertukaran yang menurunkan biaya total. Output algoritma berupa label kluster dan identitas medoid untuk tiap kluster, yang kemudian divalidasi kembali menggunakan *Silhouette Score* terhadap label final.

Visualisasi hasil klusterisasi dilakukan menggunakan *Multidimensional Scaling* (MDS) dua dimensi berbasis matriks jarak *Gower* yang sama. MDS dipilih karena bekerja langsung pada matriks jarak tanpa mengasumsikan variabel numerik kontinu dengan memproyeksikan matriks jarak ke dalam koordinat dua dimensi sambil mempertahankan kemiripan relatif antar wilayah. Kemudian, dibuat *scatter plot* dengan warna titik yang membedakan keanggotaan kluster. Profil tiap kluster dianalisis melalui statistik deskriptif berupa rerata variabel numerik dan proporsi kategori untuk merumuskan tipologi sekolah penerima bantuan.

HASIL DAN PEMBAHASAN

EKSPLORASI DATA

Untuk memahami karakteristik data penerima bantuan revitalisasi, dilakukan eksplorasi data melalui analisis statistik deskriptif sebagaimana disajikan pada Tabel 1. Hasil analisis menunjukkan bahwa nilai anggaran revitalisasi memiliki variasi yang cukup besar, dengan rentang mulai dari Rp233 juta hingga Rp8,24 miliar dan rata-rata sebesar Rp1,779 miliar. Variasi ini mengindikasikan adanya perbedaan kebutuhan dan skala revitalisasi antar sekolah penerima bantuan. Secara geografis, distribusi penerima bantuan terkonsentrasi di Pulau Jawa. Dari total 1.465 SMK pelaksana revitalisasi, sebanyak 627 sekolah berada di Pulau Jawa, diikuti Sumatera sebanyak 352 sekolah dan Sulawesi sebanyak 190 sekolah, sementara pulau lainnya masing-masing mencatat jumlah penerima di bawah 100 sekolah. Dominasi Pulau Jawa juga tercermin pada tingkat provinsi, di mana Jawa Tengah (179 SMK), Jawa Timur (177 SMK), dan Jawa Barat (173 SMK) menjadi wilayah dengan jumlah penerima bantuan terbanyak. Kondisi ini selaras dengan tingginya jumlah SMK dan kepadatan penduduk dipulau dan provinsi-provinsi tersebut dibandingkan wilayah lain di Indonesia. Pada tingkat kabupaten/kota, penerima bantuan terbanyak berada di Kabupaten Bogor (26 SMK), Kabupaten Garut (20 SMK), dan Kabupaten Malang (18 SMK). Meskipun demikian, program revitalisasi tidak hanya terfokus pada wilayah dengan konsentrasi sekolah tinggi, tetapi juga menjangkau daerah Tertinggal, Terdepan, Dan Terluar (3T) sebesar 11,40 persen. Temuan ini menunjukkan adanya upaya pemerintah untuk mendorong pemerataan pembangunan pendidikan melalui program revitalisasi SMK.

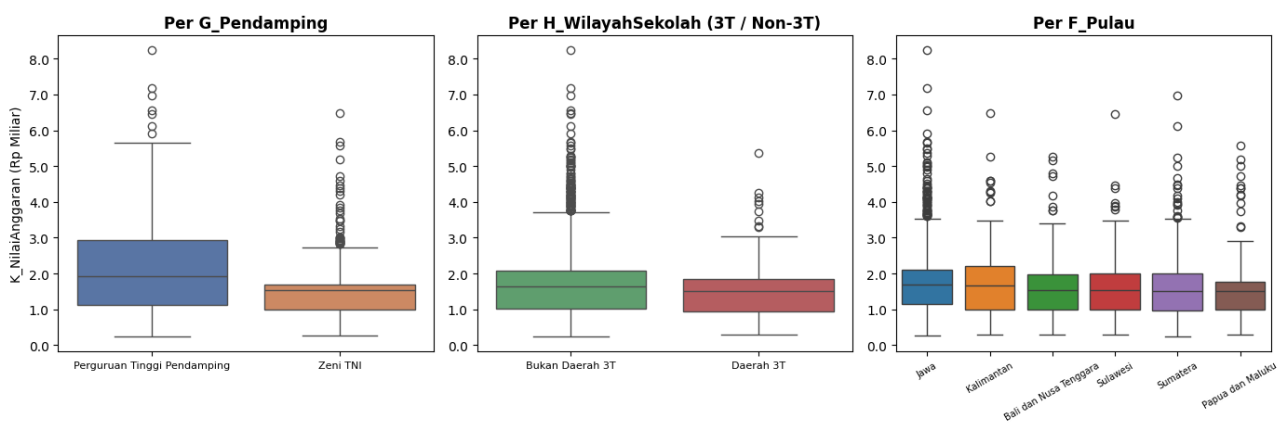
Tabel 1. Statistik Deskriptif dan Distribusi Variabel Penelitian

Variabel	Kategori/Ukuran	Nilai
A. Variabel Numerik		
Nilai Anggaran (Rp Miliar)	Minimum	0,233
	Maksimum	8,243
	Rerata	1,779
	Median	1,620
	Standar Deviasi	1,052
	Outlier (IQR)	110 (7,51%)
B. Variabel Kategorik		
Provinsi	Jawa Tengah	179 (12,22%)
	Jawa Timur	177 (12,08%)
	Jawa Barat	173 (11,81%)
	Lainnya	936 (63,89%)
Kabupaten	Bogor	26 (1,77%)

	Garut	20 (1,37%)
	Malang	18 (1,23%)
	Lainnya	1401 (95,63%)
Status Sekolah	Swasta	744 (50,78%)
	Negeri	721 (49,22%)
Pendamping	Zeni TNI	847 (57,82%)
	Perguruan Tinggi Pendamping	618 (42,18%)
Wilayah Sekolah	Bukan Daerah 3T	1.298 (88,60%)
	Daerah 3T	167 (11,40%)
Jenis Kelamin Kepsek	Laki-laki	1.116 (76,18%)
	Perempuan	349 (23,82%)
Jadwal Bimtek	Tahap 1	9 (0,61%)
	Tahap 2	123 (8,40%)
	Tahap 3	186 (12,70%)
	Tahap lainnya	1147 (78,29%)
Pernah Dapat Bantuan sebelumnya	Ya	855 (58,36%)
	Tidak	610 (41,64%)
Penerimaan dana tahap awal	Juni	181 (12,35%)
	Juli	446 (30,44%)
	Agustus	152 (10,38%)
	September	449 (30,65%)
	Oktober	237 (16,18%)
Durasi Pelaksanaan	Pelaksanaan tahun tunggal	1.202 (82,05%)
	Pelaksanaan lintas tahun	263 (17,95%)
Pulau	Jawa	627 (42,80%)
	Sumatera	352 (24,03%)
	Kalimantan	97 (6,62%)
	Papua dan Maluku	102 (6,96%)
	Sulawesi	190 (12,97%)
	Bali dan Nusa Tenggara	96 (6,55%)
	Luar Negeri	1 (0,07%)

Dilihat dari karakteristik kelembagaan sekolah, proporsi penerima bantuan antara sekolah negeri dan swasta relatif seimbang, yakni masing-masing sebesar 49,22 persen dan 50,78 persen. Komposisi ini

menunjukkan bahwa program revitalisasi dirancang secara inklusif tanpa dominasi salah satu status sekolah. Dari sisi pendampingan, mayoritas sekolah didampingi oleh Zeni TNI yang mencakup 847 SMK, sedangkan 618 sekolah lainnya memperoleh pendampingan dari Perguruan Tinggi Pendamping. Karakteristik kepemimpinan sekolah menunjukkan bahwa sebagian besar sekolah dipimpin oleh kepala sekolah laki-laki, yaitu sebanyak 1.116 SMK, sementara sekolah yang dipimpin kepala sekolah perempuan berjumlah 349 SMK. Selain itu, mayoritas penerima bantuan merupakan sekolah yang sebelumnya telah memperoleh program serupa, yaitu sebanyak 855 SMK, sedangkan penerima bantuan pertama kali berjumlah 610 SMK. Pelaksanaan bimbingan teknis bagi sekolah pelaksana dilakukan secara bertahap. Dari sisi durasi pelaksanaan, mayoritas program revitalisasi diselesaikan dalam pola pelaksanaan tahun tunggal yaitu sebanyak 1.202 sekolah atau sebesar 82,05 persen, sementara 263 sekolah lainnya mengikuti pola pelaksanaan lintas tahun. Tingginya proporsi pelaksanaan tahun tunggal ini menunjukkan bahwa implementasi program revitalisasi SMK secara nasional berjalan relatif efektif meskipun dilaksanakan pada skala yang sangat besar dan melibatkan karakteristik penerima bantuan yang beragam.

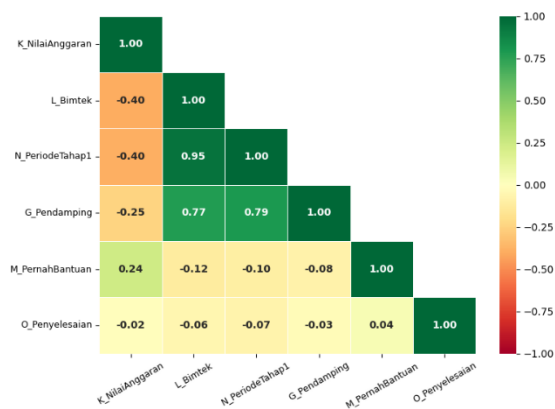


Gambar 1. Distribusi Anggaran per kategori

Sebelum tahap klusterisasi, dilakukan analisis korelasi untuk memahami pola keterkaitan antar variabel. Variabel numerik dan biner dianalisis menggunakan koefisien Spearman, sedangkan asosiasi antar variabel kategorik dianalisis menggunakan Cramér's V. Pendekatan terdiferensiasi ini dipilih karena lebih sesuai dengan tipe data masing-masing variabel. Gambar 1 menunjukkan bahwa kelompok Perguruan Tinggi Pendamping memiliki median dan sebaran interkuartil yang lebih tinggi dibandingkan Zeni TNI yang menandakan bahwa perguruan tinggi pendamping lebih banyak mendampingi SMK dengan nilai anggaran yang lebih besar dibandingkan dengan Zeni TNI. Pada dimensi wilayah, median anggaran kelompok non-3T sedikit lebih tinggi dengan sebaran lebih lebar, sementara kelompok 3T cenderung lebih homogen. Dari sisi antarpulau, Jawa menunjukkan sebaran interkuartil paling lebar, sedangkan Papua dan Maluku serta Sumatera memiliki sebaran lebih sempit dengan median lebih rendah,

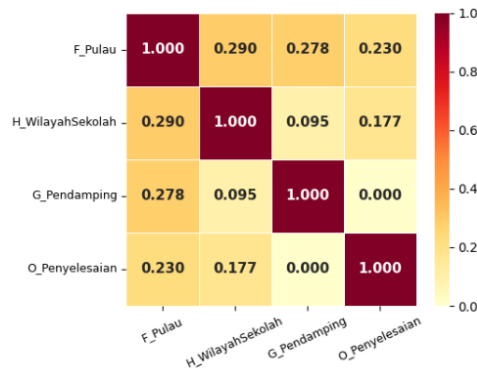
mengindikasikan bahwa variasi alokasi anggaran lebih banyak ditentukan oleh lokasi pulau dibandingkan status keterpencilan wilayah.

Analisis korelasi Spearman pada Gambar 2 menunjukkan adanya hubungan terbalik antara nilai anggaran dengan periode penyaluran dana tahap awal ($r = -0,40$), yang mengindikasikan bahwa sekolah dengan anggaran lebih besar cenderung menerima pencairan dana lebih awal. Variabel lain yang dianalisis Spearman umumnya menunjukkan korelasi lemah, mengindikasikan tidak adanya redundansi kuat antar variabel numerik dan biner pada dataset.



Gambar 2. Analisis Spearman

Analisis Cramér's V yang diilustrasikan pada Gambar 3 menunjukkan bahwa variabel Pulau memiliki asosiasi positif dengan ketiga variabel program lainnya, yaitu Wilayah Sekolah (0,290), Pendamping (0,278), dan Durasi Pelaksanaan (0,230). Pola ini mengindikasikan bahwa lokasi geografis pulau menjadi faktor diferensiator penting dalam pelaksanaan program, di mana distribusi skema pendamping, proporsi wilayah 3T, dan pola durasi pelaksanaan bervariasi secara sistematis antar pulau. Selain itu ditemukan juga bahwa durasi pelaksanaan menunjukkan korelasi sangat lemah dengan seluruh variabel lain pada analisis Spearman ($|r| < 0,07$) maupun Cramér's V (0,000 dengan Pendamping). Hal ini mengindikasikan bahwa tidak ada faktor tunggal yang dominan dalam menjelaskan pola durasi pelaksanaan program, sehingga pendekatan klasterisasi multi-dimensi diperlukan untuk menangkap kombinasi faktor yang berkontribusi terhadap karakteristik program.



Gambar 2. Analisis Cramer's

NORMALISASI DATA

Proses normalisasi diterapkan pada variabel numerik utama, yaitu nilai anggaran dengan tujuan menyamakan skala agar variabel anggaran memiliki kontribusi proporsional dalam perhitungan jarak (*Gower distance*) pada tahap klusterisasi. Hasil normalisasi memetakan seluruh nilai anggaran ke dalam rentang [0,1] (dari 0,000 hingga 1,000). Skala yang seragam ini membuat variabel anggaran yang semula memiliki rentang miliaran rupiah tidak akan mendominasi jarak *Gower* dibandingkan kontribusi variabel kategorik.

MATRIKS JARAK GOWER

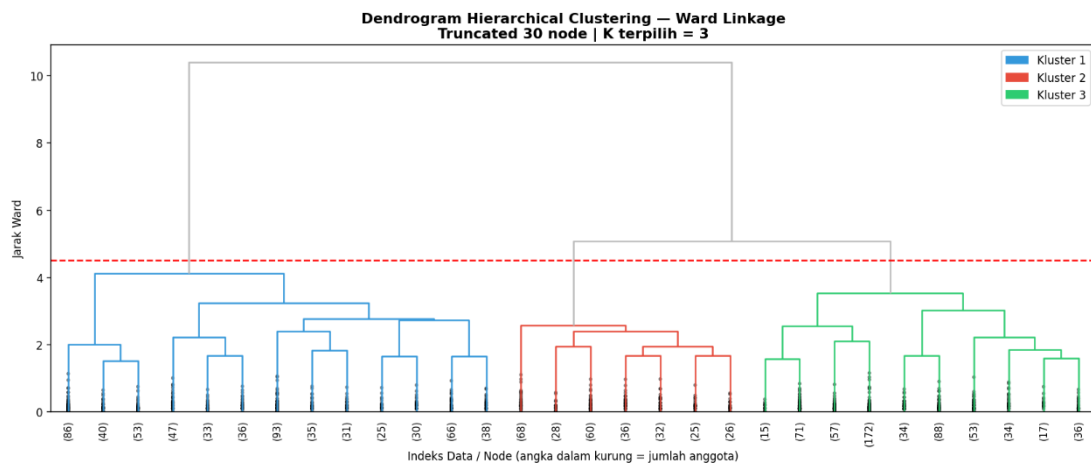
Langkah awal untuk klusterisasi yakni dengan membuat matriks jarak. Pada penelitian ini dipilih matriks Jarak *Gower* (*Gower distance*) karena data penelitian bersifat campuran (*mixed-type*), yaitu terdiri atas 11 variabel kategorikal (provinsi, kabupaten, status sekolah, pulau, pendamping, wilayah sekolah, jenis kelamin kepek, pengalaman mendapat bantuan, periode pencairan tahap awal, periode bimtek (bimbingan teknis), dan durasi pelaksanaan) serta 1 variabel numerik yang (Nilai Anggaran). *Gower Distance* menghitung jarak parsial untuk setiap variabel secara berbeda dimana variabel kategorikal menggunakan *simple matching coefficient* (bernilai 0 jika sama dan 1 jika berbeda), sedangkan variabel numerik menggunakan selisih absolut yang dinormalisasi. Pendekatan ini memungkinkan semua variabel diperlakukan langsung tanpa perlu transformasi *dummy* variable pada kolom kategorikal seperti Provinsi dan Pulau. Hasil perhitungan menghasilkan matriks jarak simetris berukuran 1.465×1.465 . Nilai jarak minimum antar sekolah yang berbeda sebesar 0,0083 dan nilai jarak maksimum sebesar 0,9682. Rentang ini mencerminkan adanya variasi profil yang cukup besar antar sekolah penerima bantuan dalam dataset.

PENENTUAN JUMLAH KLASTER OPTIMAL

Penentuan jumlah klaster optimal dilakukan melalui kombinasi pendekatan visual menggunakan *Hierarchical Clustering* dan evaluasi kuantitatif menggunakan *Silhouette Score* karena tidak ada satu metode tunggal yang dianggap superior untuk seluruh jenis data (Gagolewski et al., 2021). Kombinasi dendrogram *Ward Linkage* dan *Silhouette Score* merupakan pendekatan yang paling banyak digunakan dan tervalidasi dalam klusterisasi *mixed-type* data, baik untuk dataset kesehatan, sosial, maupun pendidikan (Szepannek et al., 2025).

Hierarchical Clustering

Hierarchical Clustering dengan metode *Ward linkage* dipilih karena dapat menunjukkan proses pengelompokan data melalui dendrogram tanpa perlu menentukan jumlah klaster sejak awal. Selain itu, metode Ward menghasilkan klaster yang lebih kompak dan relatif seimbang dengan cara meminimalkan perbedaan atau variasi di dalam setiap klaster (Großwendt et al., 2019). Dendrogram yang dihasilkan disajikan pada Gambar 1 menunjukkan bahwa struktur hierarki pemisahan yang jelas terlihat pada rentang $K = 2$ hingga $K = 4$, di mana terdapat lompatan jarak yang signifikan pada level tersebut. Garis merah pada dendrogram menunjukkan posisi pemotongan untuk $K = 3$ yang dipilih sebagai jumlah klaster optimal.



Gambar 3. Dendrogram *Hierarchical Clustering* dengan Ward Linkage

Silhouette Score

Evaluasi menggunakan *Silhouette Score* digunakan karena metode ini mampu mengukur kualitas klaster berdasarkan tingkat kedekatan anggota dalam klaster yang sama (kohesi) dan perbedaan antar klaster (separasi) secara bersamaan. Selain itu, *Silhouette Score* kompatibel dengan matriks jarak *Gower* dan banyak direkomendasikan untuk mengevaluasi hasil klusterisasi pada data campuran (Shutaywi & Kachouie, 2021). Evaluasi kuantitatif terhadap jumlah klaster optimal dilakukan menggunakan *Silhouette*

Score untuk $K = 2$ hingga $K = 6$. *Silhouette Score* mengukur seberapa baik setiap objek berada dalam klasternya sendiri dibandingkan klaster terdekat, dengan rentang nilai -1 hingga 1 di mana nilai mendekati 1 mengindikasikan pengelompokan yang lebih baik. Hasil evaluasi disajikan pada Tabel 2.

Tabel 2. *Silhouette Score* untuk Setiap Nilai K

Jumlah Klaster (K)	<i>Silhouette Score</i>	Keterangan
2	0,2501	Tertinggi
3	0,1861	Dipilih
4	0,1422	—
5	0,1048	—
6	0,1116	—

Silhouette Score tertinggi diperoleh pada $K = 2$ (0,2501), diikuti $K = 3$ (0,1861) dengan selisih 0,064. Meskipun secara kuantitatif $K = 2$ unggul tipis, penelitian ini memilih $K = 3$ sebagai jumlah klaster optimal berdasarkan empat pertimbangan. Pertama, pendekatan dendrogram *Ward Linkage* (Gambar 4) memperlihatkan tiga lompatan jarak yang jelas, mengisyaratkan struktur hierarki tiga kelompok yang stabil secara visual. Kedua, inspeksi terhadap solusi $K = 2$ menunjukkan bahwa pemisahan utamanya hanya mereplikasi dikotomi skema pendamping (Zeni TNI vs Perguruan Tinggi), suatu informasi yang sudah dapat diperoleh dari variabel tunggal sehingga tidak memberikan nilai analitis tambahan dibandingkan tabulasi silang biasa. Sebaliknya, solusi $K = 3$ berhasil memisahkan dimensi pengalaman institusional (penerima berulang vs penerima pertama kali) yang merupakan kontribusi substantif bagi perumusan kebijakan terdiferensiasi. Ketiga, nilai *Silhouette* 0,1861 untuk data campuran berdimensi tinggi dengan satu variabel numerik dan sebelas variabel kategorik masih berada dalam rentang yang dapat diterima sesuai literatur klusterisasi data campuran, mengingat sebagian besar kontribusi jarak berasal dari variabel kategorik biner-multilevel yang secara inheren menghasilkan nilai *Silhouette* lebih rendah dibandingkan data numerik murni (Gagolewski et al., 2021; Szepannek et al., 2025). Keempat, dari perspektif kebijakan, tiga klaster memberikan tipologi yang lebih informatif untuk merancang skema pendampingan terdiferensiasi dibandingkan dua klaster yang terlalu kasar.

HASIL CLUSTERING K-MEDOIDS FASTERPAM

Aplikasi algoritma K-Medoids *FasterPAM* dengan $K = 3$ menghasilkan distribusi sebagaimana disajikan pada Tabel 3. Klaster 2 merupakan kelompok terbesar dengan 600 sekolah (41,0%), diikuti Klaster 1 sebanyak 454 sekolah (31,0%), dan Klaster 3 sebanyak 411 sekolah (28,0%). Distribusi antar

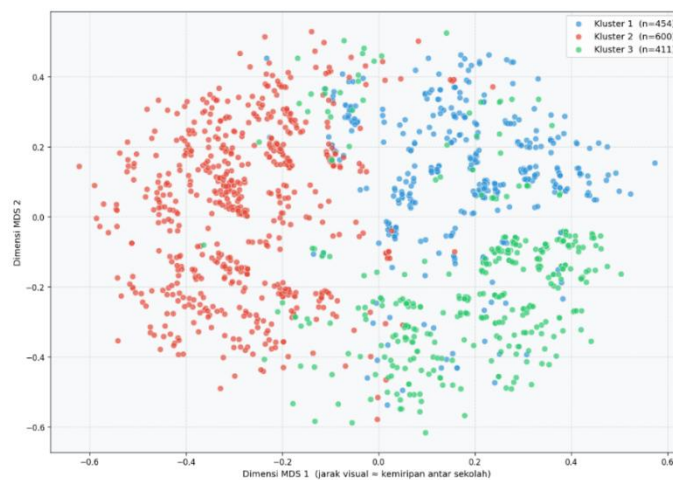
klaster yang relatif seimbang ini mengindikasikan bahwa algoritma berhasil membentuk pengelompokan yang proporsional tanpa dominasi ekstrem pada satu klaster.

Tabel 3. Distribusi Sekolah per Klaster Hasil K-Medoids

Klaster	Jumlah Sekolah	Persentase (%)	Warna
1	454	31,0	Biru
2	600	41,0	Merah
3	411	28,0	Hijau
Total	1.465	100,0	—

VISUALISASI MDS DUA DIMENSI

Scatter plot proyeksi MDS dua dimensi disajikan pada Gambar 5. Pada *scatter plot*, setiap titik merepresentasikan satu sekolah penerima bantuan dan warna titik menunjukkan keanggotaan klaster hasil K-Medoids. Jarak visual antar titik mengaproksimasi kemiripan profil antar sekolah dimana semakin berdekatan posisi dua titik, semakin mirip karakteristik keduanya. Dari visualisasi tampak bahwa ketiga klaster menempati wilayah yang dapat dibedakan pada ruang MDS, meskipun terdapat tumpang tindih pada batas antar klaster, terutama di area tengah grafik. Hal ini merupakan kondisi yang wajar mengingat data asli berdimensi 12 yang diproyeksikan ke dalam ruang dua dimensi. Klaster 2 (merah, n=600) merupakan klaster terbesar dan mendominasi sisi kiri bidang MDS (nilai MDS 1 negatif) dengan sebaran yang relatif luas dan merata. Klaster 1 (biru, n=454) cenderung mengelompok di kuadran kanan atas, mencerminkan kesamaan karakteristik antar sekolah dalam kelompok ini. Sementara itu, Klaster 3 (hijau, n=411) lebih banyak terkonsentrasi di kuadran kanan bawah, meskipun sebagian anggotanya tersebar di area tengah dan bercampur dengan kedua klaster lainnya.



Gambar 3. *Scatter Plot MDS 2D*

Kemudian, analisis profil dilakukan dengan menghitung rata-rata setiap variabel numerik dan proporsi kategori pada setiap kluster. Ringkasan profil disajikan pada Tabel 4.

Tabel 4. Profil Rata-rata Karakteristik per Kluster

Karakteristik	Klaster 1	Klaster 2	Klaster 3
Distribusi Anggota			
Jumlah sekolah (n)	454	600	411
Persentase total	31,0%	41,0%	28,0%
Karakteristik Finansial			
Rata-rata Nilai Anggaran (ternormalisasi)	0,176	0,237	0,148
Rata-rata Nilai Anggaran (Rp Miliar)	1,641	2,130	1,421
Karakteristik Temporal-Operasional			
Periode Bimtek	Periode 7	Periode 4	Periode 8
Periode penerimaan dana tahap awal	September	Juli	September
Karakteristik Institusional			
Status Sekolah Negeri (%)	34,4%	79,8%	20,9%
Pendamping Zeni TNI (%)	91,0%	11,7%	88,6%
Jenis Kelamin Kepsek Laki-laki (%)	79,3%	73,7%	76,4%
Karakteristik Historis-Operasional			
Pernah Dapat Bantuan (%)	94,3%	66,0%	7,5%
Proporsi Pelaksanaan lintas tahun (%)	17,4%	17,0%	20,0%
Karakteristik Geografis			
Wilayah 3T (%)	5,1%	16,3%	5,1%
Pulau dominan	Jawa	Sumatera	Jawa
Provinsi dominan	Jawa Timur, Jawa Tengah, Jawa Barat	Sulawesi Selatan, Jawa Tengah, Riau	Jawa Barat, Jawa Tengah, Sumatera Utara

Berdasarkan Tabel 4 dan Tabel 5, ketiga kluster menampilkan profil yang berbeda secara substantif. Klaster 1 (n=454, 31%) dicirikan oleh dominasi pendamping Zeni TNI (91,0%), proporsi penerima berulang tertinggi (94,3%), anggaran menengah (Rp1,641 miliar), penjadwalan bimtek terutama pada Periode 7, dan konsentrasi geografis di Pulau Jawa (65,2%) dengan proporsi pelaksanaan lintas tahun sebesar 17,4%. Klaster 2 (n=600, 41%) memiliki anggaran tertinggi (Rp2,130 miliar), proporsi wilayah 3T

tertinggi (16,3%), pendamping Zeni TNI minim (11,7%), penjadwalan bimtek terutama pada Periode 4, dominasi Sumatera (36,0%), dengan proporsi pelaksanaan lintas tahun 17,0% (terendah di antara ketiga klaster). Klaster 3 (n=411, 28%) berisi mayoritas penerima pertama kali (92,5%), pendamping Zeni TNI tinggi (88,6%), anggaran terkecil (Rp1,421 miliar), penjadwalan bimtek pada periode akhir dengan dominasi Pulau Jawa (49,4%), dan teridentifikasi sebagai kelompok dengan kebutuhan pendampingan paling intensif.

Tabel 5. Distribusi Pulau per Klaster (%)

Pulau	Klaster 1	Klaster 2	Klaster 3
Jawa	65,2%	21,3%	49,4%
Sumatera	12,3%	36,0%	19,5%
Sulawesi	6,4%	20,0%	10,0%
Kalimantan	5,1%	7,5%	7,1%
Bali-NT	4,2%	7,8%	7,3%
Papua-Maluku	6,8%	7,3%	6,6%

Distribusi anggota klaster berdasarkan pulau menunjukkan pola sebaran geografis yang khas. Klaster 1 sangat terkonsentrasi di Pulau Jawa (65,2%), konsisten dengan profilnya sebagai kelompok sekolah berpengalaman di wilayah aksesibel. Klaster 2 menunjukkan profil geografis paling beragam dengan dominasi Sumatera (36,0%) dan kontribusi tinggi dari Sulawesi (20,0%), sejalan dengan karakteristik klaster yang menjangkau wilayah 3T dengan tantangan geografis kompleks. Klaster 3 didominasi Pulau Jawa (49,4%) namun dengan distribusi yang lebih merata ke pulau-pulau lain, menandakan ekspansi program ke sekolah baru tetap mempertahankan basis di Jawa dengan jangkauan inklusif ke wilayah luar Jawa. Temuan ini memperkuat dasar empiris untuk perumusan kebijakan terdiferensiasi berbasis geografis sesuai karakteristik wilayah dominan tiap klaster.

Hasil klasterisasi K-Medoids dengan *Gower Distance* mengidentifikasi tiga tipologi sekolah penerima bantuan revitalisasi SMK tahun 2025 dengan distribusi yang relatif proporsional, yaitu 31 persen, 41 persen, dan 28 persen. Ketiga klaster tersebut merepresentasikan heterogenitas alami penerima bantuan dan terutama dibedakan oleh dimensi temporal pelaksanaan program yang ditentukan Direktorat SMK, seperti waktu pencairan dana tahap awal dan urutan pelaksanaan bimbingan teknis. Klaster 1 dapat dikategorikan sebagai “Sekolah Berpengalaman dengan Penjadwalan mayoritas pada Bimtek Periode 7 di Wilayah Aksesibel”, ditandai oleh dominasi penerima bantuan berulang (94,3%), pendampingan Zeni TNI (91,0%), pelaksanaan bimtek mayoritas pada Periode 7, serta konsentrasi sekolah di Pulau Jawa (65,2%). Klaster 2 merupakan “Sekolah dengan Anggaran Besar dan Penjadwalan Bimtek Periode 4 di Wilayah Menantang” yang memiliki rata-rata anggaran tertinggi sebesar Rp2,13 miliar, proporsi wilayah

3T sebesar 16,3 persen, pelaksanaan bimtek mayoritas pada Periode 4, dan pendampingan Zeni TNI yang relatif rendah (11,7%). Adapun Klaster 3 dikategorikan sebagai “Sekolah Baru dengan Penjadwalan Bimtek mayoritas di Periode 8”, dengan karakteristik dominan sebagai penerima pertama kali (92,5%), pendampingan Zeni TNI yang tinggi (88,6%), serta jadwal bimtek mayoritas pada Periode 8.

Temuan utama penelitian menggarisbawahi bahwa pola durasi pelaksanaan program merupakan hasil interaksi multifaktor yang kompleks. Klaster 2 yang menerima anggaran tertinggi dan menjangkau wilayah 3T menunjukkan proporsi pelaksanaan tahun tunggal yang relatif baik, sementara Klaster 3 yang didominasi penerima pertama kali teridentifikasi sebagai kelompok yang memerlukan perhatian pendampingan lebih intensif. Pola ini mengindikasikan bahwa selain besaran alokasi anggaran dan aspek geografis, faktor pengalaman institusional sekolah dalam mengelola bantuan, kapasitas absorptif kelembagaan, dan distribusi penjadwalan operasional turut berperan penting dalam mendukung keberhasilan implementasi program. Dari perspektif teoritis, kondisi tersebut dapat dijelaskan melalui kerangka kapasitas absorptif yang menempatkan pengalaman sebelumnya sebagai salah satu faktor pendukung keberhasilan implementasi kebijakan (Flyvbjerg & Gardner, 2023), kapasitas absorptif sekolah dalam mengelola bantuan dan pendampingan (Nel & Tacihagh, 2024; Su et al., 2024), serta tantangan logistik dan koordinasi di daerah dengan aksesibilitas rendah. Dalam konteks ini, Klaster 2 yang menjangkau wilayah 3T dengan anggaran terbesar menunjukkan proporsi pelaksanaan tahun tunggal yang relatif baik, sementara Klaster 3 yang didominasi sekolah penerima pertama kali teridentifikasi sebagai kelompok yang memerlukan pendampingan paling intensif. Temuan ini mengungkap bahwa pengalaman institusional sekolah dalam mengelola bantuan menjadi salah satu faktor yang patut diperhatikan dalam merancang strategi pendampingan, melengkapi pertimbangan lokasi geografis dan skema pendamping. Implikasinya, perhatian khusus dapat diarahkan kepada sekolah penerima pertama kali, terlepas dari lokasi geografisnya, terutama saat kombinasi penjadwalan bimtek periode akhir bertemu dengan kebutuhan adaptasi administratif sebagaimana terlihat pada Klaster 3.

Berdasarkan pelajaran dari pelaksanaan program revitalisasi SMK tahun 2025, terdapat beberapa implikasi kebijakan yang dapat dipertimbangkan untuk periode anggaran berikutnya. Pertama, profil tiga klaster dapat digunakan sebagai dasar identifikasi karakteristik calon penerima bantuan sekaligus penentuan skema pendampingan yang paling sesuai sejak awal penetapan program. Sekolah dengan karakteristik menyerupai Klaster 3, yaitu penerima pertama kali yang baru memulai keterlibatan dalam program revitalisasi, dapat menjadi prioritas pendampingan intensif sejak tahap awal untuk mendukung kelancaran implementasi. Kedua, distribusi jadwal pencairan dana dan pelaksanaan bimbingan teknis dapat dirancang lebih merata agar beban kerja terdistribusi secara proporsional; secara khusus, sekolah penerima pertama kali idealnya diberi ruang waktu adaptasi yang lebih awal dalam siklus bimtek sehingga kombinasi antara penjadwalan dan kebutuhan adaptasi administratif dapat dikelola dengan lebih baik.

Ketiga, pengembangan sistem pemantauan dini berbasis indikator progres bulanan dapat melengkapi sistem monitoring yang sudah berjalan, terutama untuk mendukung sekolah dengan profil yang memerlukan perhatian lebih, sebagaimana karakteristik Klaster 3. Keempat, praktik baik dari sekolah-sekolah pada Klaster 2 yang menunjukkan proporsi pelaksanaan tahun tunggal relatif baik dalam menjangkau wilayah 3T dengan tantangan geografis kompleks dapat didokumentasikan dan direplikasi, khususnya terkait koordinasi penjadwalan bimtek pada periode lebih awal.

KESIMPULAN

Penelitian ini memetakan profil 1.465 sekolah penerima bantuan revitalisasi SMK tahun 2025 ke dalam tiga klaster berbeda. Klaster 1 berisi 454 sekolah (31%) sebagai kelompok berpengalaman yang terkonsentrasi di Pulau Jawa, Klaster 2 merupakan kelompok terbesar berisi 600 sekolah (41%) dengan alokasi anggaran tertinggi yang menjangkau wilayah 3T, serta Klaster 3 berisi 411 sekolah (28%) sebagai penerima pertama kali. Karakteristik khas masing-masing klaster mengungkap bahwa Klaster 2 yang mendapat anggaran rata-rata terbesar di wilayah 3T menunjukkan proporsi pelaksanaan tahun tunggal yang relatif baik, sedangkan Klaster 3 yang didominasi penerima pertama kali teridentifikasi sebagai kelompok yang memerlukan pendampingan paling intensif. Temuan ini menegaskan bahwa pola durasi pelaksanaan program revitalisasi SMK didukung oleh kombinasi faktor yang mencakup pengalaman institusional sekolah, kapasitas absorptif, distribusi penjadwalan yang merata, dan alokasi anggaran yang memadai. Mengingat sekolah penerima bantuan pada periode anggaran selanjutnya akan berbeda dengan penerima 2025, pelajaran dari pelaksanaan program 2025 dapat menjadi dasar pertimbangan bagi Direktorat SMK dalam merumuskan kebijakan revitalisasi tahun mendatang yang mencakup (a) penetapan kriteria seleksi dan profil calon penerima sebelum penetapan, dengan perhatian khusus pada sekolah penerima pertama kali sebagai kelompok prioritas pendampingan, (b) pendampingan intensif sejak tahap awal bagi sekolah pemula tanpa memandang lokasi geografis, mengingat pengalaman institusional teridentifikasi sebagai salah satu faktor pendukung penting, serta (c) distribusi penjadwalan bimtek dan pencairan dana yang lebih merata sehingga sekolah pemula memperoleh ruang adaptasi yang memadai. Penelitian lanjutan dapat memperkaya kajian ini melalui peninjauan metode klusterisasi alternatif seperti K-Prototypes atau *Hierarchical Density-Based Clustering* untuk menguji *robustness* hasil pengelompokan.

DATA AVAILABLE

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

FUNDING

This work did not received any funding.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

AUTHOR CONTRIBUTION

- AWK : Writing -Review & Editing, and Supervision;
GE : Writing -Original Draft, Methodology,
TIWB : Editing, and Visualization;:
TA : Writing -Review & Editing, Validation, and Supervision.
AA : Methodology, Validation,
S : Conceptualization
MS : Formal analysis, and Writing -Review & Editing,

ACKNOWLEDGEMENT

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada Direktorat Sekolah Menengah Kejuruan, Direktorat Jenderal Pendidikan Vokasi, Kementerian Pendidikan Dasar dan Menengah atas kesempatan, dukungan data, dan fasilitasi yang diberikan selama penelitian ini berlangsung. Penulis juga menyampaikan apresiasi kepada Subdirektorat Sarana, Prasarana, dan Tata Kelola yang telah menyediakan akses terhadap database penerima bantuan revitalisasi SMK tahun anggaran 2025 sebagai sumber data utama dalam penelitian ini.

DAFTAR PUSTAKA

- Erda, G., Usdika, R. K., Pitri, R., & Erda, Z. (2023). Implementation Of The K-Medoids Method In Clustering Human Development Indexes In Indonesia. *Parameter: Journal of Statistics*, 3(2), 61–67.
- Flyvbjerg, B., & Gardner, D. (2023). *How Big Things Get Done: The Surprising Factors That Determine the Fate of Every Project, from Home Renovations to Space Exploration and Everything In Between*. Crown. <https://books.google.co.id/books?id=kmyTEAAAQBAJ>
- Gagolewski, M., Bartoszuk, M., & Cena, A. (2021). Are cluster validity measures (in) valid? *Information Sciences*, 581, 620–636. <https://doi.org/10.1016/j.ins.2021.10.004>
- Gower, J. C. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics*, 27(4), 857. <https://doi.org/10.2307/2528823>

- Großwendt, A., Röglin, H., & Schmidt, M. (2019). Analysis of Ward's Method. In *Proceedings of the Thirtieth Annual ACM-SLAM Symposium on Discrete Algorithms* (pp. 2939–2957). Society for Industrial and Applied Mathematics. <https://doi.org/10.1137/1.9781611975482.182>
- Han, J., Pei, J., & Tong, H. (2022). *Data Mining: Concepts and Techniques*.
- Liu, P., Yuan, H., Ning, Y., Chakraborty, B., Liu, N., & Peres, M. A. (2024). A modified and weighted Gower distance-based clustering analysis for mixed type data: a simulation and empirical analyses. *BMC Medical Research Methodology*, 24(1), 305. <https://doi.org/10.1186/s12874-024-02427-8>
- Nel, D., & Taeihagh, A. (2024). The soft underbelly of complexity science adoption in policymaking: towards addressing frequently overlooked non-technical challenges. *Policy Sciences*, 57(2), 403–436. <https://doi.org/10.1007/s11077-024-09531-y>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Schubert, E., & Rousseeuw, P. J. (2021). Fast and eager ϵ -medoids clustering: runtime improvement of the PAM, CLARA, and CLARANS algorithms. *Information Systems*, 101, 101804. <https://doi.org/10.1016/j.is.2021.101804>
- Shutaywi, M., & Kachouie, N. N. (2021). Silhouette Analysis for Performance Evaluation in Machine Learning with Applications to Clustering. *Entropy*, 23(6), 759. <https://doi.org/10.3390/e23060759>
- Su, Y., Bullock, H. L., Trisolini, M., & Emmons, K. M. (2024). Editorial: Learning for action in policy implementation. *Frontiers in Health Services*, 4. <https://doi.org/10.3389/frhs.2024.1515478>
- Szepannek, G., Aschenbruck, R., & Wilhelm, A. (2025). Clustering large mixed-type data with ordinal variables. *Advances in Data Analysis and Classification*, 19(3), 749–767. <https://doi.org/10.1007/s11634-024-00595-5>
- Vidyananta, I. P. A. V., & Dermawan, K. T. D. (2025). Perbandingan algoritma k-means dan k-medoids dalam pengelompokan provinsi di indonesia berdasarkan indikator keadaan sekolah dasar. *Jurnal Informatika Dan Teknik Elektro Terapan*, 13(3S1). <https://doi.org/10.23960/jitet.v13i3S1.8145>