

ANALISIS KETIMPANGAN PENDIDIKAN DASAR BERDASARKAN INFRASTRUKTUR SEKOLAH DENGAN ALGORITMA K-MEANS DI KOTA BANDUNG

Deni Muhammad*, Naufal Arief Budiman, Rian Febriansyah & Hilma Mutiara Winata
UIN Sunan Gunung Djati Bandung, Indonesia
*Email: denidencok@gmail.com**

Riwayat Artikel

Diterima: 5 November 2025
Disetujui: 5 Desember 2025
Diterbitkan: 19 Desember 2025

Abstract

Primary education plays a strategic role in shaping the quality of human resources in the future. However, disparities among elementary schools in Bandung City remain evident, particularly in terms of the availability and quality of educational infrastructure such as classrooms, laboratories, libraries, and the number of teaching staff. This study aims to analyze the inequality of primary education based on school infrastructure conditions using the K-Means algorithm. The research employs a quantitative descriptive-exploratory approach utilizing secondary data obtained from Open Data Kota Bandung. The analyzed variables include facility availability, the number of teachers, and school accreditation status. The analysis process involves data cleaning, standardization, determination of the number of clusters using the Elbow Method, and grouping with the K-Means algorithm. The findings are expected to illustrate the pattern of infrastructure inequality in primary education across Bandung City objectively and serve as a basis for local government policy formulation in achieving equitable education quality.

Keywords: educational inequality, school infrastructure, K-Means, clustering, Bandung City.

Abstrak

Pendidikan dasar memiliki peran strategis dalam membentuk kualitas sumber daya manusia di masa depan. Namun, di Kota Bandung masih ditemukan ketimpangan antar sekolah dasar, terutama dalam hal ketersediaan dan kualitas infrastruktur pendidikan seperti ruang kelas, laboratorium, perpustakaan, serta jumlah tenaga pendidik. Penelitian ini bertujuan untuk menganalisis tingkat ketimpangan pendidikan dasar berdasarkan kondisi infrastruktur sekolah dasar menggunakan algoritma K-Means. Metode penelitian yang digunakan adalah pendekatan kuantitatif deskriptif eksploratif dengan sumber data sekunder yang diperoleh dari Open Data Kota Bandung dan Dapodik. Dianalisis mencakup ketersediaan fasilitas, jumlah guru, dan status akreditasi sekolah. Proses analisis dilakukan melalui tahap pembersihan data, standarisasi, penentuan jumlah kluster menggunakan Elbow Method, dan pengelompokan menggunakan algoritma K-Means. Hasil penelitian diharapkan dapat menggambarkan pola ketimpangan infrastruktur pendidikan dasar di Kota Bandung secara objektif, serta menjadi dasar bagi pemerintah daerah dalam merumuskan kebijakan pemerataan mutu pendidikan.

Kata Kunci: Clustering, Ketimpangan Pendidikan, Infrastruktur Sekolah, K-Means, Kota Bandung.

A. PENDAHULUAN

Pendidikan dasar merupakan fondasi utama dalam membentuk kualitas sumber daya manusia dan kemajuan suatu bangsa. Melalui pendidikan dasar yang merata dan berkualitas,

setiap anak memiliki kesempatan yang sama untuk mengembangkan potensi dirinya (Fadil et al., 2023). Namun, pemerataan pendidikan di Indonesia masih menjadi tantangan besar, terutama dalam konteks infrastruktur dan sarana prasarana sekolah. Kota Bandung sebagai salah satu pusat Pendidikan di Jawa Barat mencerminkan fenomena ini dengan jelas. Di satu sisi, terdapat sekolah dasar yang memiliki fasilitas lengkap dan tenaga pendidik memadai, tetapi di sisi lain, masih banyak sekolah yang mengalami keterbatasan fasilitas, seperti ruang kelas yang tidak layak, minimnya laboratorium dan perpustakaan, serta rasio guru-siswa yang tidak seimbang (Marwan et al., 2024).

Fenomena ini menunjukkan adanya ketimpangan pendidikan yang bersifat struktural dan sistemik. Ketimpangan tersebut tidak hanya mempengaruhi kualitas proses belajar-mengajar, tetapi juga berdampak pada pencapaian hasil belajar siswa (Yusuf & Hanif, 2025). Sekolah dengan infrastruktur memadai cenderung menghasilkan prestasi akademik lebih tinggi karena dukungan fasilitas pembelajaran yang optimal, sedangkan sekolah dengan keterbatasan infrastruktur mengalami hambatan dalam mencapai mutu pendidikan yang diharapkan (Satria et al., 2025). Kondisi ini menjadi tantangan serius bagi pemerintah daerah, karena kesenjangan pendidikan pada jenjang dasar akan berdampak jangka panjang terhadap kualitas sumber daya manusia dan tingkat kesejahteraan masyarakat.

Beberapa penelitian terdahulu yang relevan menggunakan algoritma *K-Means* telah menunjukkan potensi besar metode ini dalam menganalisis data pendidikan. Mustakim & Kamal (2021) menerapkan *K-Means clustering* untuk mengklasifikasikan mutu manajemen pendidikan menengah di Indonesia berdasarkan standar seperti fasilitas, laboratorium, dan jumlah tenaga pendidik. Hasil penelitian mereka menunjukkan bahwa metode *K-Means* efektif dalam memetakan perbedaan mutu pendidikan antar wilayah secara objektif. Sementara itu, Kartikadarma et al. (2021) menggunakan algoritma *K-Means* dalam menganalisis rasio siswa terhadap kelompok belajar di sekolah dasar. Penelitian ini menegaskan kemampuan *K-Means* dalam membantu pengambilan keputusan terkait distribusi sumber daya pendidikan. Selain itu, Nurrachman (2021) membandingkan algoritma *K-Means* dan *Fuzzy C-Means* dalam pengelompokan hasil ujian nasional SMP, yang memperlihatkan bahwa *K-Means* memiliki kinerja yang efisien dalam mengidentifikasi pola pada data pendidikan.

Tabel 1. Research Gap

Judul Penelitian	Fokus	Metode	Data	Konteks
Analisis Perbandingan <i>K-Means</i> dan <i>Fuzzy C-Means</i> Hasil Ujian Nasional SMP (Nurrachman, 2021)	Menerapkan algoritma <i>K-Means</i> untuk pengelompokan hasil ujian nasional SMP	Kuantitatif eksperimental	Data sekunder hasil Ujian Nasional tahun 2019 dari Pusat Asesmen Pendidikan (Puspendik)	Penelitian dilakukan pada tingkat provinsi di seluruh Indonesia dengan fokus pada manajemen mutu pendidikan menengah.

ARTIKEL

Penerapan K-Means Clustering untuk Klasifikasi Mutu Manajemen Pendidikan Menengah di Indonesia (Mustakim & Kamal, 2021)	Mengelompokkan provinsi di Indonesia berdasarkan standar mutu manajemen pendidikan menengah di Indonesia (fasilitas, guru, laboratorium).	K-Means Clustering.	Data sekunder tahun 2020–2021 dari Kementerian Pendidikan dan Kebudayaan (Kemendikbud).	Penelitian berfokus pada skala nasional dengan konteks peningkatan efisiensi pembagian rombongan belajar di sekolah dasar.
Penerapan Data Mining dalam Mendukung Keputusan Rasio Siswa dan Rombongan Belajar di Sekolah Dasar di Indonesia Menggunakan Metode K-Means (Kartikadarma et al., 2021)	Menganalisis rasio siswa terhadap kelompok belajar di sekolah dasar.	K-Means Clustering.	Data sekunder tahun 2020 dari <i>Data Pokok Pendidikan (Dapodik)</i> Kemdikbud.	Penelitian berfokus pada analisis hasil ujian nasional SMP di Indonesia untuk membandingkan performa algoritma <i>K-Means</i> dan <i>Fuzzy C-Means</i> .

Sumber: Diolah Peneliti, 2025

Penelitian ini memiliki keunikan yang membedakannya dari penelitian terdahulu. Secara kontekstual, penelitian ini berfokus secara spesifik pada ketimpangan infrastruktur pendidikan dasar di Kota Bandung, bukan dalam lingkup nasional atau pada jenjang pendidikan menengah sebagaimana dilakukan oleh penelitian sebelumnya. Dari sisi metodologis, penelitian ini tetap menggunakan pendekatan kuantitatif berbasis algoritma *K-Means clustering*, namun penerapannya diarahkan untuk mengidentifikasi pola kesenjangan infrastruktur antar sekolah dasar berdasarkan kesamaan karakteristik fasilitas pendidikan, jumlah guru, serta status akreditasi.

Selain itu, penelitian ini memiliki nilai empiris yang kuat karena menggunakan data aktual yang bersumber dari *Open Data Kota Bandung* dan Dapodik, sehingga hasil analisis yang diperoleh bersifat faktual dan dapat diverifikasi. Hasil yang diharapkan tidak hanya menggambarkan kondisi ketimpangan infrastruktur pendidikan dasar secara objektif, tetapi juga dapat digunakan sebagai dasar pengambilan keputusan dalam perencanaan pemerataan mutu pendidikan di tingkat daerah. Oleh karena itu, penelitian ini memiliki kontribusi ganda, yaitu sebagai pengembangan kajian akademis dalam penerapan algoritma *K-Means* di bidang pendidikan serta sebagai instrumen praktis dalam mendukung kebijakan berbasis bukti (*evidence-based policy making*) di sektor pendidikan dasar.

Urgensi dari penelitian ini terletak pada pentingnya memahami ketimpangan pendidikan dasar secara empiris melalui pendekatan analisis data yang komprehensif. Pemerintah daerah memerlukan peta ketimpangan infrastruktur yang akurat untuk menentukan prioritas intervensi kebijakan pendidikan, seperti peningkatan fasilitas di sekolah yang tertinggal dan pemerataan distribusi tenaga pendidik. Selain itu, hasil penelitian ini diharapkan dapat menjadi dasar bagi pengembangan sistem perencanaan berbasis data di sektor pendidikan, yang pada akhirnya mendukung terwujudnya pemerataan mutu pendidikan dasar di Kota Bandung.

Dalam menghadapi permasalahan tersebut, dibutuhkan pendekatan berbasis data untuk memahami dan memetakan ketimpangan infrastruktur pendidikan secara objektif. Salah satu metode yang dapat digunakan adalah analisis clustering dengan algoritma K-Means. Metode ini mampu mengelompokkan sekolah-sekolah berdasarkan kemiripan kondisi infrastrukturnya, sehingga dapat diidentifikasi kelompok sekolah dengan fasilitas tinggi, sedang, maupun rendah. Melalui pendekatan ini, hasil analisis dapat memberikan gambaran empiris tentang sebaran ketimpangan infrastruktur pendidikan dasar di Kota Bandung dan menjadi dasar dalam perumusan kebijakan pemerataan pendidikan yang lebih efektif dan berkeadilan.

Dengan demikian, penelitian ini tidak hanya memiliki nilai akademis dalam penerapan algoritma K-Means untuk analisis ketimpangan pendidikan, tetapi juga nilai praktis dalam mendukung pengambilan keputusan berbasis data bagi pemerintah daerah. Hasil penelitian diharapkan dapat membantu mengidentifikasi wilayah prioritas yang membutuhkan intervensi kebijakan guna meningkatkan pemerataan kualitas pendidikan dasar di Kota Bandung.

B. KAJIAN PUSTAKA

Ketimpangan Pendidikan

Ketimpangan pendidikan merupakan kondisi di mana terdapat perbedaan yang signifikan dalam akses, kualitas, dan sarana pendidikan antar individu, kelompok, atau wilayah (Edo & Yasin, 2024). Ketimpangan ini muncul akibat distribusi sumber daya pendidikan yang tidak merata, baik dalam hal tenaga pendidik, fasilitas belajar, maupun dukungan infrastruktur (Aprilia et al., 2024). Menurut John Rawls (1971) dalam *A Theory of Justice*, keadilan sosial dapat tercapai apabila setiap individu memiliki kesempatan yang sama untuk mengakses sumber daya penting, termasuk pendidikan (Arianto et al., 2025). Ketimpangan dalam pendidikan menunjukkan bahwa prinsip keadilan distributif tersebut belum terwujud secara optimal.

Dalam konteks pendidikan dasar di Kota Bandung, ketimpangan pendidikan tampak nyata dari perbedaan kondisi antar sekolah, di mana sebagian sekolah memiliki fasilitas lengkap dan modern, sedangkan sebagian lainnya kekurangan sarana pembelajaran yang memadai (Maharani et al., 2024). Ketimpangan semacam ini tidak hanya berdampak pada kualitas proses belajar-mengajar, tetapi juga memperlebar jurang pencapaian akademik antar sekolah (Wijayanti et al., 2024). Dengan demikian, teori ketimpangan pendidikan menjadi dasar untuk memahami bagaimana ketidakseimbangan distribusi infrastruktur sekolah dapat memengaruhi mutu pendidikan dan pemerataan kesempatan belajar.

Pemerataan Pendidikan

Pemerataan pendidikan memiliki tujuan untuk menciptakan sistem pendidikan yang memberikan kesempatan belajar yang sama bagi seluruh peserta didik tanpa diskriminasi sosial, ekonomi, maupun geografis (Vega et al., 2024). Menurut Tilaar (2000), pemerataan pendidikan bukan sekadar memastikan bahwa setiap anak dapat bersekolah, tetapi juga menjamin bahwa mereka mendapatkan layanan pendidikan yang setara dari segi kualitas. Hakim (2016) menegaskan bahwa pemerataan pendidikan harus mencakup dimensi aksesibilitas, kualitas, dan relevansi agar seluruh peserta didik memperoleh pengalaman belajar yang adil dan layak.

Dalam konteks penelitian ini, teori pemerataan pendidikan memberikan kerangka untuk menafsirkan hasil analisis *K-Means clustering* dalam memetakan kelompok sekolah berdasarkan kondisi infrastrukturnya. Sekolah-sekolah yang tergolong dalam kelompok dengan fasilitas rendah menggambarkan area yang perlu mendapatkan perhatian lebih dalam kebijakan pemerataan pendidikan. Dengan demikian, teori ini tidak hanya berfungsi sebagai landasan konseptual, tetapi juga menjadi instrumen analisis untuk memahami sebaran ketimpangan infrastruktur dan arah kebijakan yang perlu ditempuh oleh pemerintah daerah.

Analisis Clustering dengan Algoritma K-Means

Analisis clustering merupakan salah satu metode dalam *unsupervised learning* yang digunakan untuk mengelompokkan sekumpulan data berdasarkan kemiripan karakteristik tanpa memerlukan label awal (Arvi, 2024). Han, Kamber, & Pei (2012) dalam *Data Mining: Concepts and Techniques* menjelaskan bahwa clustering bertujuan untuk meminimalkan variasi antar objek dalam satu kelompok (*intra-cluster distance*) dan memaksimalkan perbedaan antar kelompok (*inter-cluster distance*). Algoritma *K-Means* menjadi salah satu metode clustering yang paling populer karena sederhana, efisien, dan mampu menangani data dalam skala besar (Sari, 2024). Proses *K-Means* dimulai dengan menentukan jumlah kluster (k), kemudian menghitung jarak antara setiap data dengan pusat cluster (*centroid*), dan memperbarui posisi centroid hingga mencapai kestabilan atau konvergensi.

Penerapan algoritma *K-Means* dalam konteks pendidikan telah terbukti efektif dalam berbagai penelitian, seperti yang dilakukan oleh Mustakim & Kamal (2021), Kartikadarma et al. (2021), dan Nurrachman (2021). Metode ini memungkinkan peneliti untuk mengidentifikasi pola-pola tersembunyi dalam data pendidikan, misalnya dalam hal rasio siswa-guru, distribusi fasilitas, atau hasil belajar. Dalam penelitian ini, algoritma *K-Means* digunakan untuk mengelompokkan sekolah dasar di Kota Bandung berdasarkan kondisi infrastruktur seperti jumlah ruang kelas, laboratorium, perpustakaan, jumlah guru, dan status akreditasi. Melalui hasil pengelompokan ini, dapat diketahui sekolah-sekolah mana yang berada pada kelompok fasilitas tinggi, sedang, dan rendah.

Secara teoretis, penerapan algoritma *K-Means* dalam konteks penelitian ini memiliki dua manfaat utama. Pertama, memberikan gambaran empiris tentang tingkat ketimpangan infrastruktur pendidikan dasar di Kota Bandung secara objektif dan berbasis data. Kedua, menyediakan dasar yang kuat bagi pengambil kebijakan untuk merumuskan strategi pemerataan pembangunan pendidikan yang lebih efektif dan berkeadilan. Dengan demikian, teori-teori yang dikemukakan dalam kajian pustaka ini menjadi pijakan penting dalam menjelaskan fenomena ketimpangan pendidikan serta mendukung interpretasi hasil penelitian secara ilmiah dan terarah.

C. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode analisis *clustering* menggunakan algoritma *K-Means*. Tujuan dari pendekatan ini adalah untuk mengidentifikasi pola ketimpangan infrastruktur pendidikan dasar di Kota Bandung secara objektif dan terukur melalui analisis data numerik. Penelitian ini bersifat deskriptif karena berfokus pada penggambaran kondisi eksisting infrastruktur sekolah dasar tanpa melakukan intervensi langsung terhadap variabel yang diteliti. Data yang digunakan merupakan data sekunder tahun 2023 yang diperoleh dari Open Data Kota Bandung dan Data Pokok Pendidikan (Dapodik). Data ini dipilih karena bersifat aktual dan mencerminkan kondisi nyata infrastruktur pendidikan dasar di Kota Bandung. Penelitian ini menggunakan 4 indikator utama untuk mengukur kondisi infrastruktur pendidikan dasar, yaitu:

- Ketersediaan Fasilitas
- Jumlah Guru
- Jumlah Ruangan
- Akreditasi Sekolah

Tahapan analisis dilakukan dengan menstandarkan seluruh data agar memiliki skala yang sebanding, kemudian menentukan jumlah cluster optimal menggunakan Metode Elbow, dan diperkuat oleh Silhouette Score, selanjutnya menerapkan algoritma *K-Means* untuk mengelompokkan sekolah dasar berdasarkan kemiripan kondisi infrastrukturnya. Analisis

dilakukan dengan bantuan perangkat lunak Python, menggunakan pustaka Pandas dan Scikit-learn untuk pengolahan data serta Matplotlib untuk visualisasi hasil.

Dengan metode ini, penelitian diharapkan dapat memberikan gambaran empiris mengenai pengelompokan sekolah dasar di Kota Bandung berdasarkan kondisi infrastruktur yang dimiliki, sehingga membantu pemerintah daerah dalam merumuskan kebijakan pemerataan pendidikan yang lebih terarah.

D. HASIL DAN PEMBAHASAN

Tahapan Pengumpulan Data

Tahapan pengumpulan data dalam penelitian ini dilakukan secara sistematis untuk memperoleh informasi yang relevan dan akurat mengenai kondisi infrastruktur pendidikan dasar di Kota Bandung. Sumber utama data berasal dari Open Data Kota Bandung dan Data Pokok Pendidikan (Dapodik). Kedua sumber ini dipilih karena memiliki kredibilitas tinggi dan menyediakan data pendidikan yang terverifikasi secara resmi oleh pemerintah.

Proses pengumpulan data dimulai dengan identifikasi variabel penelitian yang berkaitan dengan kondisi infrastruktur sekolah dasar. Variabel tersebut meliputi jumlah ruang kelas, ketersediaan fasilitas laboratorium dan perpustakaan, jumlah guru, serta status akreditasi sekolah. Data ini kemudian dikumpulkan dari basis data daring yang tersedia secara publik, khususnya melalui portal data.bandung.go.id dan laman dapodik.kemdikbud.go.id. Pemilihan variabel dilakukan berdasarkan relevansinya terhadap tujuan penelitian, yaitu untuk menganalisis ketimpangan pendidikan dasar melalui perbedaan kondisi infrastruktur antar sekolah.

Selanjutnya, data yang diperoleh melalui portal tersebut diunduh dalam format spreadsheet (.csv) agar dapat diolah menggunakan perangkat lunak analisis data. Setelah itu dilakukan proses pembersihan data (data cleaning), yaitu tahap untuk memastikan keakuratan dan konsistensi data. Pada tahap ini, entri yang memiliki nilai kosong, data duplikat, atau tidak sesuai dengan identitas sekolah di Kota Bandung dihapus atau disesuaikan dengan referensi resmi dari Dapodik. Langkah ini penting agar hasil analisis tidak bias akibat kesalahan input atau ketidaksesuaian data.

Selain pembersihan data, dilakukan juga standarisasi data (data normalization) untuk menyamakan skala antar variabel. Hal ini penting karena setiap variabel memiliki satuan yang berbeda misalnya jumlah guru dihitung dalam orang, sedangkan akreditasi menggunakan kategori huruf (A, B, C). Normalisasi dilakukan agar seluruh variabel dapat dianalisis secara proporsional menggunakan algoritma K-Means, yang sensitif terhadap perbedaan skala data.

Setelah seluruh proses pengolahan awal selesai, data siap digunakan untuk tahap analisis kluster. Hasil dari tahapan pengumpulan data ini menghasilkan satu set data terintegrasi yang menggambarkan profil infrastruktur sekolah dasar di Kota Bandung secara komprehensif, sehingga dapat digunakan sebagai dasar untuk mengidentifikasi pola ketimpangan dan perbedaan karakteristik antar sekolah.

Tahap Klasterisasi

Tahap klasterisasi merupakan inti dari proses analisis dalam penelitian ini, di mana data yang telah dikumpulkan dan diproses sebelumnya dikelompokkan berdasarkan tingkat kemiripan karakteristik infrastrukturnya menggunakan algoritma K-Means. Tujuan utama dari tahap ini adalah untuk mengidentifikasi pola ketimpangan infrastruktur pendidikan dasar di Kota Bandung secara empiris dan terukur, sehingga dapat diketahui kelompok sekolah dengan kondisi fasilitas yang relatif tinggi, sedang, dan rendah.

Proses klasterisasi diawali dengan menentukan jumlah kluster optimal (k) yang akan digunakan dalam analisis. Penentuan nilai k dilakukan menggunakan Metode Elbow, yaitu dengan menghitung nilai Within-Cluster Sum of Squares (WCSS) untuk beberapa

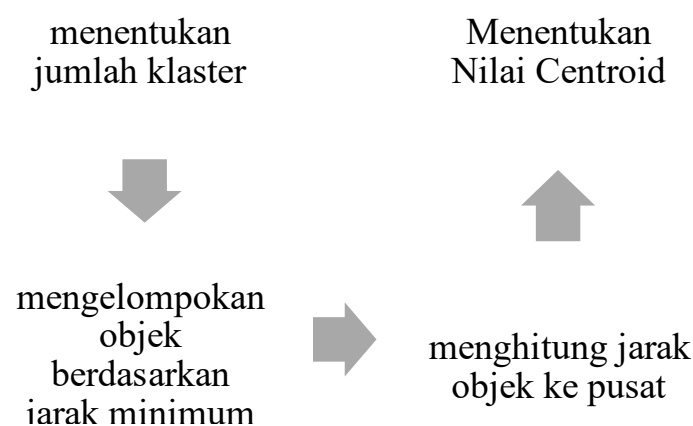
kemungkinan jumlah klaster. Grafik hasil perhitungan menunjukkan titik “siku” (elbow) di mana penurunan nilai WCSS mulai melambat secara signifikan. Titik tersebut dianggap sebagai jumlah klaster paling efisien karena menyeimbangkan antara kompleksitas model dan tingkat akurasi pengelompokan.

Setelah jumlah klaster optimal ditentukan, algoritma K-Means mulai dijalankan dengan membagi seluruh data sekolah dasar ke dalam kelompok berdasarkan jarak terdekat antara data dengan titik pusat klaster (centroid). Setiap centroid mewakili karakteristik rata-rata dari sekolah dalam klaster tersebut. Proses ini berlangsung secara iteratif — posisi centroid diperbarui terus-menerus hingga perubahan posisi antar iterasi tidak lagi signifikan, yang menandakan bahwa model telah mencapai kondisi konvergen dan hasil klaster stabil.

Hasil klasterisasi kemudian divisualisasikan untuk mempermudah interpretasi. Sekolah dasar dikelompokkan menjadi tiga kategori besar: klaster dengan infrastruktur tinggi, klaster dengan infrastruktur sedang, dan klaster dengan infrastruktur rendah. Sekolah dalam klaster pertama umumnya memiliki jumlah ruang kelas yang memadai, fasilitas laboratorium dan perpustakaan yang lengkap, jumlah guru proporsional, serta status akreditasi tinggi (A). Sementara itu, sekolah pada klaster kedua berada pada kondisi menengah, di mana sebagian fasilitas tersedia namun belum optimal. Adapun sekolah dalam klaster ketiga menunjukkan keterbatasan fasilitas fisik dan sumber daya manusia, dengan status akreditasi yang cenderung lebih rendah.

Dari hasil klasterisasi ini terlihat adanya ketimpangan yang cukup jelas antar sekolah dasar di Kota Bandung. Perbedaan paling menonjol terdapat pada ketersediaan ruang kelas dan fasilitas pembelajaran, yang menunjukkan bahwa infrastruktur pendidikan belum tersebar merata di seluruh wilayah kota. Klaster dengan fasilitas rendah umumnya ditemukan di wilayah dengan kepadatan penduduk tinggi dan keterbatasan lahan, sedangkan klaster dengan fasilitas tinggi cenderung berada di wilayah pusat kota atau daerah yang memiliki dukungan anggaran pendidikan lebih besar.

Temuan ini memberikan bukti empiris bahwa ketimpangan pendidikan dasar di Kota Bandung memiliki korelasi yang kuat dengan distribusi infrastruktur dan sumber daya pendidikan. Hasil klasterisasi menggunakan algoritma K-Means ini diharapkan dapat menjadi dasar pertimbangan bagi pemerintah daerah dalam merumuskan kebijakan pembangunan pendidikan yang lebih merata, serta dalam menentukan prioritas intervensi peningkatan fasilitas sekolah di wilayah yang tertinggal. Berikut ini akan ditampilkan urutan dari algoritma K-Means Klasterisasi yang akan dilakukan dalam penelitian ini yang akan ditunjukkan pada gambar 1.1

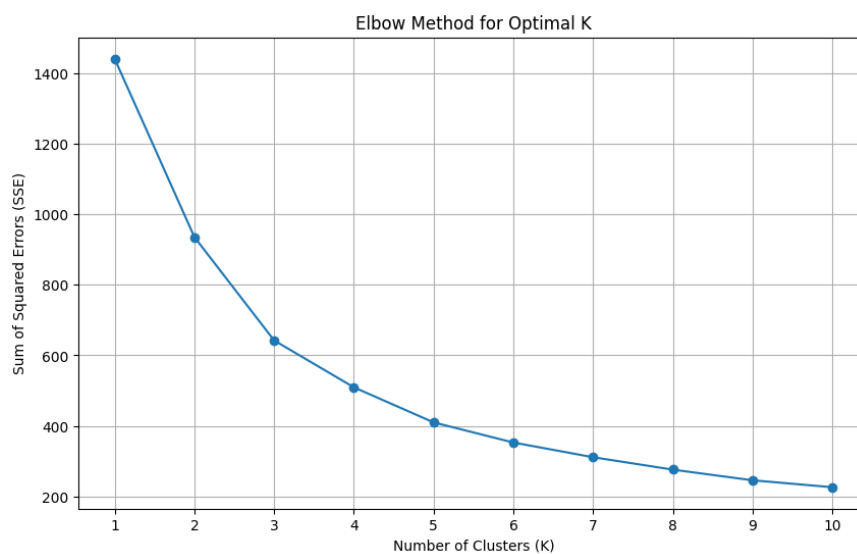


Gambar 1. Algoritma K-Means Klastering
Sumber: Diolah Peneliti, 2025

Menentukan Jumlah Kluster

Metode Elbow merupakan salah satu pendekatan paling umum untuk menentukan jumlah kluster optimal dalam analisis K-Means clustering. Prinsip utamanya adalah menilai hubungan antara jumlah kluster (K) dan jumlah kuadrat kesalahan dalam kluster (Sum of Squared Errors/SSE).

Dalam konteks penelitian ini, setiap penambahan nilai K akan menurunkan nilai SSE karena semakin banyak kluster berarti jarak antara data dan pusat klasternya semakin kecil. Namun, penurunan tersebut tidak selalu signifikan setelah titik tertentu. Titik di mana penurunan SSE mulai melandai disebut sebagai “titik siku” (elbow point), yang menandakan jumlah kluster optimal.



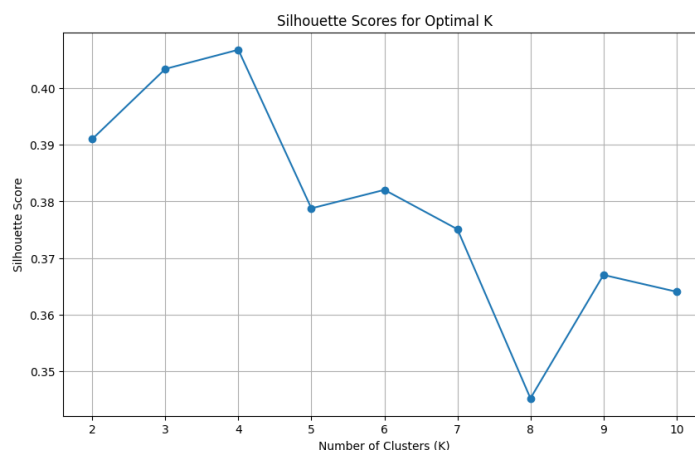
Gambar 2. Grafik Elbow
Sumber: Diolah Peneliti, 2025

Berdasarkan grafik Elbow dalam gambar 2. yang dihasilkan, terjadi penurunan SSE yang tajam dari $K=1$ hingga $K=3$, kemudian kurva mulai melandai pada $K=4$. Hal ini menunjukkan bahwa penambahan jumlah kluster setelah empat tidak lagi memberikan peningkatan yang signifikan dalam kualitas pemisahan data. Dengan demikian, $K=4$ ditetapkan sebagai jumlah kluster optimal karena pada titik tersebut keseimbangan antara kompleksitas model dan efektivitas segmentasi telah tercapai.

Secara substantif, hal ini mengindikasikan bahwa sekolah dasar di Kota Bandung dapat dikelompokkan menjadi empat kategori berbeda berdasarkan kombinasi tingkat fasilitas, jumlah ruang kelas, jumlah guru, dan akreditasi yang dimiliki.

Metode Silhouette Score digunakan untuk mengevaluasi kualitas pembentukan kluster yang dihasilkan. Nilai Silhouette mengukur seberapa mirip suatu objek dengan klasternya sendiri dibandingkan dengan kluster lain. Nilainya berkisar antara -1 hingga 1, di mana:

- Nilai mendekati 1 menunjukkan bahwa objek sangat sesuai dengan klasternya dan jauh dari kluster lain (pemisahan baik).
- Nilai mendekati 0 menunjukkan bahwa objek berada di batas antara dua kluster.
- Nilai negatif menunjukkan bahwa objek mungkin lebih cocok berada di kluster lain (pemisahan buruk).



Gambar 3. Silhouette
Sumber: Diolah Peneliti, 2025

Hasil perhitungan pada gambar 3. Silhouette pada dataset ini menunjukkan nilai tertinggi berada pada K=4, dengan skor sekitar 0,41. Nilai ini tergolong moderat, menandakan bahwa struktur kluster cukup jelas dan jarak antar kluster masih terjaga dengan baik. Meskipun tidak sempurna, nilai ini sudah memadai untuk analisis sosial seperti dalam konteks pendidikan, di mana variabilitas antar sekolah wajar terjadi karena perbedaan sumber daya, kondisi geografis, dan kebijakan lokal.

Dengan demikian, hasil metode Silhouette memperkuat temuan dari metode Elbow bahwa empat kluster (K=4) merupakan jumlah kelompok paling ideal untuk menggambarkan pola kesamaan antar sekolah dasar di Kota Bandung.

Menentukan Kelompok Data

Berdasarkan hasil klusterisasi menggunakan algoritma K-Means dengan K=4, data sekolah dasar di Kota Bandung terbagi menjadi empat kelompok utama.

bps_nama_kabupaten_kota	npsn	nama_sekolah
0 KOTA BANDUNG	20245064	SDN 003 PAGARSIH KOTA BANDUNG
1 KOTA BANDUNG	20245706	SDN 015 KRESNA KOTA BANDUNG
2 KOTA BANDUNG	20245530	SDN 004 CISARANTEN KULON KOTA BANDUNG
3 KOTA BANDUNG	20245212	SDN 002 KARANGMULYA KOTA BANDUNG
4 KOTA BANDUNG	20244836	SDN 014 CIGONDEWAH KOTA BANDUNG

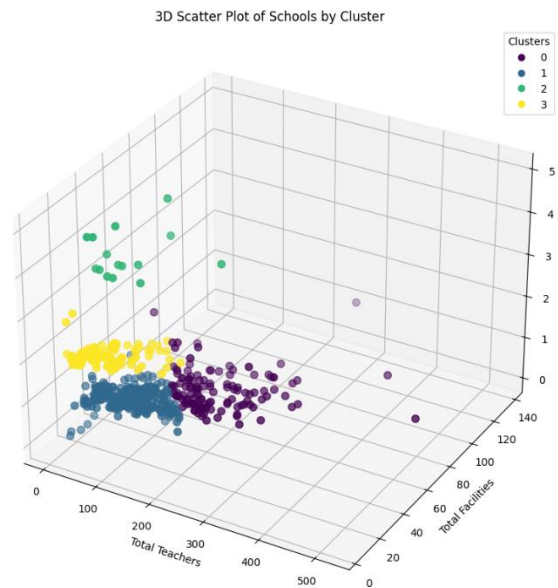
status_sekolah	akreditasi	semester_ajaran	tahun_ajaran	tahun	total_guru
0 NEGERI	A	2	2022/2023	2023	180.0
1 NEGERI	A	2	2022/2023	2023	188.0
2 NEGERI	A	2	2022/2023	2023	136.0
3 NEGERI	A	2	2022/2023	2023	119.0
4 NEGERI	A	2	2022/2023	2023	145.0

total_fasilitas	cluster
0 41.0	0
1 31.0	0
2 19.0	1
3 19.0	1
4 17.0	1

Gambar 4. Hasil Pengelompokan
Sumber: Diolah Peneliti, 2025

Gambar 4. memperlihatkan sebagian data hasil pengelompokan yang menampilkan variabel utama seperti total guru, total fasilitas, dan akreditasi sekolah beserta label klasternya. Dari tabel tersebut terlihat bahwa sekolah dengan jumlah guru dan fasilitas lebih tinggi umumnya masuk ke Cluster 0, sedangkan sekolah dengan sumber daya lebih terbatas cenderung masuk ke Cluster 1 dan seterusnya. Hal ini menunjukkan bahwa proses klusterisasi

berhasil mengelompokkan sekolah berdasarkan kemiripan karakteristik sumber daya yang dimilikinya.



Gambar 5. 3D Scater Plot
Sumber: Diolah Peneliti, 2025

Sementara itu, Pada gambar 5. menampilkan 3D Scatter Plot hasil klasterisasi yang menggambarkan distribusi sekolah berdasarkan tiga variabel utama: total guru (sumbu X), total fasilitas (sumbu Y), dan akreditasi sekolah (sumbu Z). Setiap warna merepresentasikan klaster yang berbeda, sehingga pola pemisahan antarklaster dapat terlihat jelas. Sekolah dengan nilai fasilitas dan jumlah guru tinggi terkonsentrasi pada satu kelompok, sedangkan sekolah dengan nilai indikator rendah membentuk kelompok tersendiri. Visualisasi ini memperkuat bukti bahwa pembentukan empat klaster sudah tepat, karena data membentuk empat wilayah yang relatif terpisah dan konsisten dengan karakteristik masing-masing kelompok.

Keempat warna pada grafik 3D menunjukkan bahwa proses klasterisasi berhasil memetakan sekolah ke dalam kelompok yang memiliki perbedaan jelas dalam hal kapasitas sumber daya.

- Warna ungu menandakan kelompok unggul,
- biru menunjukkan kelompok cukup kuat,
- hijau menggambarkan kelompok berkembang, dan
- kuning mewakili kelompok terbatas atau perlu intervensi kebijakan.

Evaluasi Kinerja Clustering

Evaluasi kinerja dilakukan untuk menilai seberapa baik proses klasterisasi yang telah dibentuk oleh algoritma K-Means dalam mengelompokkan sekolah dasar di Kota Bandung. Pada tahap ini, dua ukuran utama digunakan sebagai dasar penilaian, yaitu Sum of Squared Errors (SSE) dan Silhouette Score. Kedua indikator ini digunakan untuk memastikan bahwa jumlah klaster yang dipilih tidak hanya efisien secara matematis, tetapi juga mampu menggambarkan struktur data secara konsisten.

Tabel 2. Jumlah Kluster

Jumlah Kluster (K)	Sum of Squared Errors (SSE)	Silhouette Score	Observasi/Interpretasi
1	1440.0	Tidak Terdefinisi	Titik awal untuk SSE.
2	935.0	0.391	Penurunan SSE yang tajam. Silhouette Score mulai terdefinisi.
3	642.2	0.403	Penurunan SSE yang signifikan. Skor Silhouette meningkat.
4	509.4	0.407	Titik 'siku' yang jelas pada SSE. Skor Silhouette tertinggi, menunjukkan kluster optimal.
5	410.0	0.379	Penurunan SSE melambat. Skor Silhouette menurun.
6	353.0	0.382	Penurunan SSE sedikit meningkat tetapi tidak melebihi K=4.
7	311.4	0.375	Penurunan SSE minimal. Skor Silhouette menurun kembali.
8	276.2	0.345	Penurunan SSE minimal. Skor Silhouette terendah.
9	246.0	0.367	Penurunan SSE sangat kecil. Skor Silhouette sedikit membaik.
10	226.4	0.364	Penurunan SSE minimal. Skor Silhouette tidak menunjukkan perbaikan signifikan.

Sumber: Diolah Peneliti, 2025

Berdasarkan hasil yang ditunjukkan pada table 2. evaluasi, nilai Sum of Squared Errors (SSE) mengalami penurunan yang sangat tajam dari kluster pertama (K=1) hingga kluster ketiga (K=3). Hal ini menunjukkan bahwa semakin banyak jumlah kluster, semakin kecil jarak antara setiap data dengan pusat klasternya. Namun, mulai dari K=4, penurunan nilai SSE tidak lagi signifikan dan cenderung melandai. Kondisi tersebut menunjukkan munculnya titik “siku” atau elbow point pada K=4, yang menandakan jumlah kluster optimal telah tercapai. Pada titik ini, efisiensi model sudah baik karena penambahan kluster berikutnya tidak memberikan peningkatan yang berarti dalam penyebaran data.

Sementara itu, hasil Silhouette Score menunjukkan kecenderungan yang sejalan. Nilai Silhouette mulai terbentuk dan meningkat dari K=2 hingga mencapai nilai tertinggi pada K=4, yaitu sekitar 0,407. Nilai ini termasuk kategori cukup baik, menandakan bahwa struktur kluster sudah terbentuk dengan jelas dan jarak antar kluster cukup terpisah. Setelah K=4, nilai Silhouette justru menurun, yang berarti penambahan jumlah kluster membuat batas antar kelompok menjadi semakin kabur dan mengurangi konsistensi internal data dalam setiap kluster.

Dari kedua hasil tersebut dapat disimpulkan bahwa jumlah kluster terbaik untuk menggambarkan variasi sekolah dasar di Kota Bandung adalah empat kluster (K=4). Pada titik ini, keseimbangan antara efisiensi pengelompokan dan kejelasan batas antar kelompok telah tercapai. Dengan demikian, hasil klusterisasi ini dianggap paling optimal dalam menggambarkan perbedaan karakteristik antar sekolah berdasarkan empat indikator utama yang digunakan, yaitu jumlah guru, jumlah ruang kelas, ketersediaan fasilitas, dan akreditasi sekolah.

Menentukan Nilai Centroid

Tahap ini bertujuan untuk menentukan pusat kluster (centroid) yang menjadi representasi rata-rata karakteristik dari setiap kelompok sekolah. Nilai centroid menggambarkan posisi sentral dari variabel penelitian yakni jumlah guru, jumlah fasilitas, dan akreditasi sekolah—setelah proses iterasi K-Means mencapai kondisi konvergen (stabil). Dengan demikian, centroid menjadi indikator utama untuk mengenali profil umum dari masing-masing kluster yang terbentuk.

Tabel 3. Nilai Centroid

--- Nilai Centroid (Skala Asli) ---			
	total_guru	total_fasilitas	akreditasi_encoded
0	208.211382	45.073171	1.056911
1	98.677221	23.626914	0.951807
2	92.666667	25.533333	4.266667
3	71.193548	18.881720	2.021505

Sumber: Diolah Peneliti, 2025

Berdasarkan hasil perhitungan tabel 3. diperoleh empat centroid dengan nilai pada skala asli data. Kluster 0 memiliki rata-rata jumlah guru sebanyak 208 orang dan fasilitas sekitar 45

unit, sedangkan klaster 1 memiliki rata-rata 98 guru dan 23 fasilitas. Klaster 2 menunjukkan nilai akreditasi tertinggi dengan skor 4,27, meskipun jumlah guru dan fasilitasnya relatif sedang, yaitu sekitar 93 guru dan 25 fasilitas. Sementara itu, klaster 3 menjadi kelompok dengan sumber daya paling rendah—memiliki rata-rata hanya 71 guru dan 19 fasilitas, serta akreditasi sekitar 2,02.

Hasil tersebut menunjukkan bahwa setiap klaster memiliki ciri khas yang berbeda. Klaster 0 menggambarkan sekolah dengan sumber daya manusia dan sarana yang tinggi; klaster 1 merepresentasikan sekolah menengah dengan mutu stabil; klaster 2 menandakan sekolah dengan kinerja akademik tinggi meskipun sumber dayanya terbatas; dan klaster 3 mengindikasikan sekolah dengan keterbatasan paling besar di semua aspek.

Untuk memperkuat pemahaman terhadap posisi setiap klaster, dilakukan analisis deskriptif tambahan berupa mean, median, dan standar deviasi per klaster. Nilai median digunakan untuk melihat posisi nilai tengah dari masing-masing indikator agar tidak terpengaruh oleh data ekstrem.

Tabel 4. Mean Per Cluster

--- Mean per Cluster ---			
	akreditasi_encoded	total_guru	total_fasilitas
cluster			
0	1.056911	208.211382	45.073171
1	0.951807	98.677221	23.626914
2	4.266667	92.666667	25.533333
3	2.021505	71.193548	18.881720
--- Median per Cluster ---			
	akreditasi_encoded	total_guru	total_fasilitas
cluster			
0	1.0	198.0	43.0
1	1.0	100.0	23.0
2	4.0	85.0	21.0
3	2.0	62.0	18.0
--- Standard Deviation per Cluster ---			
	akreditasi_encoded	total_guru	total_fasilitas
cluster			
0	0.265528	65.378000	14.861200
1	0.214605	41.536132	7.114191
2	0.457738	41.709312	14.730273
3	0.145848	38.298955	7.822556

Sumber: Diolah Peneliti, 2025

Hasilnya menunjukkan bahwa median jumlah guru dan fasilitas cenderung mendekati nilai rata-rata, menandakan distribusi data yang relatif seimbang di setiap kelompok.

Selain itu, nilai standard deviation menunjukkan tingkat keragaman di dalam klaster. Klaster 0 memiliki penyebaran tertinggi pada jumlah guru (sekitar 65 guru), menandakan bahwa meskipun termasuk kelompok besar, variasi ukuran antar sekolahnya masih cukup lebar. Sebaliknya, klaster 3 memiliki standar deviasi paling kecil (38 guru dan 7 fasilitas), yang berarti sekolah-sekolah dalam kelompok ini cenderung homogen dengan karakteristik serupa, yakni sumber daya terbatas. Klaster 2 menunjukkan variasi cukup tinggi dalam akreditasi, menandakan adanya perbedaan mutu meskipun berada dalam kelompok yang sama.

Secara keseluruhan, hasil ini memperlihatkan bahwa nilai centroid tidak hanya berfungsi sebagai pusat perhitungan klaster, tetapi juga sebagai representasi statistik dari kondisi nyata tiap kelompok sekolah. Dengan adanya empat centroid yang berbeda secara signifikan, dapat disimpulkan bahwa model K-Means dengan K=4 berhasil membentuk struktur pengelompokan yang logis dan relevan dengan kondisi pendidikan dasar di Kota Bandung.

Menghitung Jarak Objek Ke Pusat

Tahap ini bertujuan untuk mengetahui kedekatan setiap objek (sekolah) terhadap pusat klaster (centroid) yang telah terbentuk. Jarak ini menjadi ukuran seberapa mirip karakteristik suatu sekolah dengan profil umum kelompoknya. Semakin kecil jarak yang dimiliki suatu sekolah terhadap centroid-nya, semakin besar tingkat keseragaman karakter sekolah tersebut

dengan klaster yang menaunginya. Sebaliknya, jarak yang lebih besar menunjukkan adanya perbedaan karakter atau potensi penyimpangan relatif terhadap klaster tempat sekolah tersebut bergabung.

Tabel 5. *Assigned Centroid*

```

--- df_merged with Distance to Assigned Centroid ---
id kode_provinsi nama_provinsi bps_kode_kabupaten_kota \
0 1 32 JAWA BARAT 3273
1 2 32 JAWA BARAT 3273
2 3 32 JAWA BARAT 3273
3 4 32 JAWA BARAT 3273
4 5 32 JAWA BARAT 3273

bps_nama_kabupaten_kota npsn nama_sekolah \
0 KOTA BANDUNG 20245064 SDN 003 PAGARSIH KOTA BANDUNG
1 KOTA BANDUNG 20245706 SDN 015 KRESNA KOTA BANDUNG
2 KOTA BANDUNG 20245530 SDN 004 CISARANTEN KULON KOTA BANDUNG
3 KOTA BANDUNG 20245212 SDN 002 KARANGMULYA KOTA BANDUNG
4 KOTA BANDUNG 20244836 SDN 014 CIGONDEWAH KOTA BANDUNG

status_sekolah akreditasi semester_ajaran tahun_ajaran tahun total_guru \
0 NEGERI A 2 2022/2023 2023 188.0
1 NEGERI A 2 2022/2023 2023 188.0
2 NEGERI A 2 2022/2023 2023 136.0
3 NEGERI A 2 2022/2023 2023 119.0
4 NEGERI A 2 2022/2023 2023 145.0

```

Sumber: Diolah Peneliti, 2025

Tabel 5. menampilkan sebagian data sekolah dasar di Kota Bandung yang telah digabungkan dengan hasil perhitungan jarak ke pusat klaster. Setiap sekolah diidentifikasi berdasarkan kode NPSN dan nama sekolah, kemudian ditempatkan dalam klaster yang paling sesuai berdasarkan hasil algoritma K-Means.

Kolom cluster menunjukkan kelompok tempat sekolah tergabung, sedangkan kolom distance_to_centroid menampilkan nilai jarak Euclidean antara sekolah tersebut dengan pusat klasternya. Sebagai contoh, sekolah seperti SDN 003 Pagarsih dan SDN 015 Kresna memiliki jarak sekitar 0,49–1,03, menandakan bahwa sekolah tersebut cukup dekat dengan pusat klaster dan memiliki karakteristik yang sangat mirip dengan profil rata-rata kelompoknya. Sementara nilai jarak yang lebih besar, misalnya di atas 1, mengindikasikan bahwa sekolah tersebut masih berada dalam klaster yang sama, namun memiliki perbedaan kecil dalam hal jumlah guru, fasilitas, atau nilai akreditasi.

Secara keseluruhan, data ini memperlihatkan hasil konkret dari proses iterasi K-Means, di mana setiap sekolah telah menemukan posisi paling dekat terhadap salah satu dari empat centroid yang terbentuk.

Tabel 6. Fasilitas Kluster

```

total_fasilitas cluster distance_to_centroid
0 41.0 0 0.497316
1 31.0 0 1.035454
2 19.0 1 0.623167
3 19.0 1 0.439581
4 17.0 1 0.806282

--- Jarak Objek Ke Pusat (Summary per Klaster) ---
count mean std min 25% 50% 75% \
cluster
0 123.0 1.183600 0.822049 0.125139 0.772599 1.003151 1.356499
1 249.0 0.725072 0.397830 0.086502 0.429848 0.680886 0.930232
2 15.0 1.153608 0.642882 0.392082 0.810411 0.848740 1.615803
3 93.0 0.695489 0.386378 0.075741 0.435222 0.621963 0.894232

max
cluster
0 6.529136
1 2.242516
2 2.302262
3 1.790391

```

Sumber: Diolah Peneliti, 2025

Tabel 6. menampilkan ringkasan statistik jarak objek ke pusat untuk masing-masing klaster. Ringkasan ini mencakup jumlah sekolah dalam klaster (count), nilai rata-rata jarak (mean), sebaran data (std), serta nilai minimum dan maksimum jarak (min–max). Melalui tabel tersebut, dapat dilihat variasi tingkat kedekatan antar sekolah di dalam klaster masing-masing.

Klaster 0 memiliki rata-rata jarak terbesar yaitu 1,18, menandakan bahwa meskipun kelompok ini berisi sekolah dengan jumlah guru dan fasilitas tinggi, penyebarannya cukup bervariasi ada sekolah yang sangat dekat dengan pusat, namun ada pula yang lebih jauh. Hal ini menggambarkan keragaman karakter internal di antara sekolah-sekolah besar di Kota Bandung. Sebaliknya, klaster 3 memiliki rata-rata jarak terkecil (0,69) dengan deviasi standar rendah, menunjukkan bahwa sekolah-sekolah dalam kelompok ini lebih homogen dan memiliki karakteristik yang sangat mirip satu sama lain, baik dari sisi jumlah guru maupun fasilitas.

Nilai maksimum jarak pada setiap klaster juga memberikan informasi penting. Misalnya, jarak maksimum di klaster 0 mencapai 6,52, yang menandakan adanya satu atau dua sekolah dengan karakteristik cukup berbeda (kemungkinan outlier) dibandingkan rata-rata klaster. Namun, secara keseluruhan jarak rata-rata yang kecil pada sebagian besar klaster menunjukkan bahwa hasil pengelompokan sudah cukup stabil dan kompak.

Dari kedua tabel tersebut dapat disimpulkan bahwa hasil penghitungan jarak menunjukkan konsistensi yang baik antara anggota klaster dan pusatnya. Sekolah-sekolah dengan nilai jarak kecil memperlihatkan tingkat kemiripan yang tinggi dengan karakter umum klaster, sedangkan beberapa sekolah dengan nilai jarak lebih besar menunjukkan adanya variasi internal yang wajar.

Secara keseluruhan, tahap ini mengonfirmasi bahwa hasil klasterisasi K-Means yang dilakukan telah menghasilkan pengelompokan yang efektif, di mana setiap sekolah telah ditempatkan pada kelompok dengan profil yang paling mendekati karakteristiknya.

Rekomendasi Kebijakan

Rekomendasi kebijakan yang relevan dengan hasil analisis ketimpangan infrastruktur pendidikan dasar di Kota Bandung harus dilakukan secara komprehensif dan berkelanjutan. Pertama, pemerintah daerah perlu memfokuskan intervensi pada wilayah-wilayah yang menunjukkan kondisi fasilitas rendah, seperti yang teridentifikasi melalui hasil klasterisasi menggunakan algoritma K-Means. Upaya ini meliputi pembangunan dan peningkatan fasilitas fisik seperti ruang kelas, laboratorium, dan perpustakaan, serta penambahan tenaga pendidik yang berkualitas agar pemerataan akses dan mutu pendidikan dapat tercapai.

Selain itu, diperlukan pengembangan program pemerataan infrastruktur secara strategis dan terintegrasi, yang didukung oleh alokasi anggaran yang memadai dan berkelanjutan. Program ini harus melibatkan kolaborasi dengan berbagai pihak, termasuk swasta dan masyarakat setempat, untuk memastikan keberlanjutan dan keberhasilan pembangunan fasilitas pendidikan di wilayah tertinggal. Penting juga untuk meningkatkan kapasitas dan kesejahteraan tenaga pendidik melalui pelatihan, insentif, dan peningkatan kesejahteraan agar kualitas proses belajar mengajar dapat meningkat secara menyeluruh. Kebijakan ini akan membantu mengurangi ketimpangan kualitas pendidikan yang dipengaruhi oleh perbedaan sumber daya manusia di berbagai sekolah.

Penggunaan data dan teknologi dalam perencanaan dan monitoring menjadi aspek penting lainnya. Pemerintah perlu membangun sistem data berbasis digital yang mampu memantau perkembangan infrastruktur dan kualitas pendidikan secara real-time. Data ini harus digunakan sebagai dasar pengambilan keputusan yang cepat dan tepat sasaran, sehingga intervensi dapat dilakukan secara lebih efektif dan efisien. Selain itu, evaluasi dan monitoring secara berkala terhadap hasil klasterisasi dan implementasi kebijakan harus dilakukan untuk memastikan bahwa program yang dijalankan benar-benar efektif dan sesuai dengan kondisi lapangan. Penggunaan indikator seperti SSE dan Silhouette Score dapat menjadi standar dalam menilai keberhasilan kebijakan dan melakukan penyesuaian jika diperlukan.

Terakhir, partisipasi aktif dari masyarakat, orang tua, dan stakeholder terkait sangat penting dalam proses perencanaan dan pelaksanaan program pembangunan infrastruktur dan

peningkatan mutu pendidikan. Partisipasi ini akan meningkatkan keberlanjutan dan keberhasilan kebijakan pemerataan pendidikan, serta memastikan bahwa kebutuhan lokal dapat terpenuhi secara optimal. Dengan menerapkan kebijakan tersebut secara terpadu dan berkelanjutan, diharapkan ketimpangan infrastruktur dan mutu pendidikan di Kota Bandung dapat diminimalisasi, sehingga tercipta pemerataan pendidikan yang adil dan berkeadilan.

E. KESIMPULAN

Berdasarkan hasil penelitian yang menggunakan algoritma K-Means untuk mengelompokkan sekolah dasar di Kota Bandung berdasarkan infrastruktur, dapat disimpulkan bahwa terdapat ketimpangan yang cukup signifikan antar sekolah dalam hal fasilitas, jumlah guru, dan status akreditasi. Penentuan jumlah kluster optimal dilakukan dengan metode Elbow dan Silhouette Score, keduanya menunjukkan bahwa empat kluster merupakan pilihan terbaik untuk merepresentasikan variasi karakteristik sekolah di wilayah tersebut. Hasil pengelompokan ini memperlihatkan adanya sekolah dengan kondisi fasilitas yang sangat baik hingga yang membutuhkan intervensi, yang tersebar di berbagai wilayah kota.

Penggunaan visualisasi warna dan penghitungan jarak antar sekolah ke pusat kluster membantu dalam memahami tingkat homogenitas dan heterogenitas antar kelompok sekolah. Kesimpulan ini mendukung pentingnya pengambilan kebijakan berbasis data untuk pemerataan infrastruktur pendidikan, sehingga dapat meningkatkan kualitas pendidikan secara lebih adil dan merata di seluruh wilayah Kota Bandung. Dengan demikian, analisis ini menjadi dasar yang kuat untuk merancang strategi kebijakan yang lebih efektif dalam mengatasi ketimpangan dan meningkatkan mutu pendidikan dasar di tingkat daerah.

REFERENSI

- Aprilia, E., Afindi, A., & Saputra, H. (2024). Kasus Pendayagunaan Sumber Daya Pendidikan. *Jurnal Pendidikan, Bahasa Dan Budaya*, 3(4), 112–124.
- Arianto, Y. F., Agustiani, M. F., Shalzabilla, S., & Mayangsari, D. A. (2025). Konsep Keadilan Restoratif Dalam Perspektif Teori Keadilan John Rawls. *Nusantara: Jurnal Pendidikan, Seni, Sains Dan Sosial Humaniora*, 3(01).
- Arvi, R. N. (2024). *Analisis Segmentasi Pasar Berdasarkan Penjualan Produk Menggunakan Metode Clustering K-Means, K-Medoids Dan Agglomerative Hierarchical Clustering*. Universitas Komputer Indonesia.
- Edo, A., & Yasin, M. (2024). Dampak kesenjangan akses pendidikan dan faktor ekonomi keluarga terhadap mobilitas sosial. *Jurnal Ilmu Pendidikan & Sosial (SINOVA)*, 2(3), 317–326.
- Fadil, K., Amran, A., & Alfaien, N. I. (2023). Peningkatan Kualitas Pendidikan Dasar Melalui Implementasi Kurikulum Merdeka Belajar Dalam Mewujudkan Sustainable Developments Goal's. *ATTADIB*, 7(2).
<https://dapo.kemendikdasmen.go.id/>
<https://opendata.bandung.go.id/>
- Marwan, I., Gustaman, R. F., & Gandi, A. (2024). Dikotomi Sarana Dan Prasarana Pendidikan Di Indonesia. *Cendekia: Jurnal Pendidikan Dan Pemberdayaan Masyarakat*, 2(3), 155–161.
- Maharani, N., Khoirunnisa, N., & Putri, S. P. (2024). Analisis Masalah Kesenjangan Sosial Di Sekolah Dasar. *PARADIGM: Journal Of Multidisciplinary Research and Innovation*, 2(01), 1–16.
- Sari, P. (2024). Algoritma K-Means Clustering: Sebuah Studi Literatur. *Jurnal Informatika*, 1(01), 1–7.

- Vega, A., Maharani, I. V. A., Putri, J. A., Hartono, M. R. A. M., & Navridya, R. U. (2024). Kesetaraan Akses Pendidikan: Analisis Pengimplementasian Nilai Pancasila Dalam Pemerataan Akses Pendidikan Di Indonesia. *Lentera Ilmu*, 1(2), 44–57.
- Wijayanti, A., Darmawan, A. W., & Marwan, I. (2024). Isu-Isu Kontemporer Pendidikan Indonesia: Kesenjangan Pendidikan. *Cendekia: Jurnal Pendidikan Dan Pemberdayaan Masyarakat*, 2(3), 187–192.
- Yusuf, A., & Hanif, M. (2025). Dampak Desentralisasi Pendidikan terhadap Ketimpangan Akses di Indonesia. *Jurnal Inovasi Penelitian Ilmu Pendidikan Indonesia*, 153–166.