

Penerapan Metode Clustering dengan Algoritma K-Means untuk Pengelompokan Data Sekolah Menengah di Kabupaten Muna Barat

Sigrid Tamriesfatno ^{1*}, Riska Kherani ², dan Sarwono ²

¹ Institusi Kesehatan dan Teknologi Buton Raya 1; email : sigridaaris@gmail.com

² Universitas Teknologi Akba Makassar 2; email : riskakherani1@gmail.com

³ Universitas Muhammadiyah Kupang 3; email : sarwono@unmuhkupang.ac.id

* Penulis Korespondensi

Abstract: Improving the quality of education requires a comprehensive understanding of the conditions and characteristics of each educational institution. West Muna Regency has various secondary schools with diverse profiles and challenges. Until now, school grouping has often been done manually, which does not always accurately reflect the overall characteristics of the data. This study aims to cluster secondary schools in West Muna Regency using the K-Means algorithm as a clustering method to identify hidden patterns in school data, such as the number of students, teachers, staff, facilities, and location. The research method involves several stages, including data collection, method analysis, software implementation, and cluster testing. The clustering results produced three school clusters with different characteristics. Cluster 1 consists of schools with the most complete resources, Cluster 2 includes the largest number of schools with varying resources, and Cluster 3 represents schools with moderate conditions. These findings are expected to serve as a basis for formulating more targeted educational policies, ensuring equitable distribution of resources, and improving the quality of education in West Muna Regency.

Keywords: K-Means, Clustering, Data Mining, Secondary School, West Muna

Abstrak: Peningkatan mutu pendidikan memerlukan pemahaman yang komprehensif terhadap kondisi dan karakteristik masing-masing institusi pendidikan. Kabupaten Muna Barat memiliki berbagai sekolah menengah dengan profil dan tantangan yang beragam. Selama ini, pengelompokan sekolah sering dilakukan secara manual, yang tidak selalu mencerminkan karakteristik data secara menyeluruh. Penelitian ini bertujuan untuk mengelompokkan sekolah menengah di Kabupaten Muna Barat menggunakan algoritma K-Means sebagai metode clustering untuk mengidentifikasi pola tersembunyi dalam data sekolah, seperti jumlah siswa, guru, pegawai, fasilitas, dan lokasi. Metode penelitian dilakukan melalui tahapan pengumpulan data, analisis metode, implementasi perangkat lunak, serta pengujian cluster. Hasil pengelompokan menghasilkan tiga klaster sekolah dengan karakteristik berbeda. Klaster 1 berisi sekolah dengan sumber daya paling lengkap, Klaster 2 merupakan kelompok dengan jumlah sekolah terbanyak namun sumber daya bervariasi, dan Klaster 3 menunjukkan sekolah dengan kondisi sedang. Temuan ini diharapkan dapat menjadi dasar dalam penyusunan kebijakan pendidikan yang lebih tepat sasaran, pemerataan sumber daya, serta peningkatan kualitas pendidikan di Kabupaten Muna Barat.

Kata kunci: K-Means, clustering, data mining, Sekolah Menengah, Muna Barat

1. Pendahuluan

Pendidikan merupakan pilar utama pembangunan suatu daerah, dan kualitasnya sangat bergantung pada pemahaman komprehensif mengenai kondisi serta karakteristik institusi pendidikan yang ada. Sarana dan prasarana sekolah yang memadai akan menunjang

Diterima: Maret, 31 2025

Direvisi: April, 14 2025

Diterima: Mei, 07 2025

Diterbitkan: Mei, 21 2025



Hak cipta: © 2025 oleh penulis.
Diserahkan untuk kemungkinan
publikasi akses terbuka
berdasarkan syarat dan ketentuan
lisensi Creative Commons
Attribution (CC BY SA) (
<https://creativecommons.org/licenses/by-sa/4.0/>)

terselenggaranya proses belajar mengajar yang efektif, serta menjamin terselenggaranya kebijakan pendidikan secara berkelanjutan, sehingga menghasilkan keunggulan akademis (Yaro, 2023)

Kabupaten Muna Barat, sebagai salah satu wilayah di Sulawesi Tenggara, memiliki sejumlah sekolah menengah dengan berbagai profil dan tantangan. Data sekolah menengah, seperti jumlah siswa, rasio guru-siswa, fasilitas, nilai rata-rata ujian, dan lokasi geografis, merupakan sumber informasi berharga yang jika dianalisis dengan baik dapat memberikan gambaran mendalam tentang disparitas dan potensi pengembangan.

Selama ini, pengelompokan sekolah seringkali dilakukan secara manual atau berdasarkan kategori administratif yang sudah ada, yang mungkin tidak selalu mencerminkan karakteristik intrinsik dan pola tersembunyi dalam data. Pendekatan ini dapat menghambat identifikasi sekolah yang memiliki kemiripan dalam aspek-aspek tertentu, namun terdistribusi secara geografis, atau sebaliknya. Kurangnya pemahaman yang mendalam mengenai kluster-kluster sekolah ini dapat menyebabkan kebijakan pendidikan yang kurang tepat sasaran, alokasi sumber daya yang tidak optimal, serta intervensi yang tidak efektif.

Analisis cluster telah banyak digunakan di berbagai bidang ilmu pengetahuan seperti ekonomi, psikologi, kesehatan, sosial masyarakat dan kependudukan (Tri et al., 2019). Salah satu metode clustering yang paling terkenal adalah K-Means. K-Means adalah algoritma clustering non-hirarki yang mengelompokkan objek-objek berdasarkan kemiripan karakteristiknya (Musharyadi et al., 2017). Proses pengelompokan ini dilakukan secara berulang atau iteratif hingga mencapai kondisi akhir tertentu, yang kemudian menghasilkan cluster-cluster sebagai outputnya (Fakhri et al., 2021). Dengan kata lain, K-Means akan menempatkan objek yang punya karakteristik sama dalam satu cluster, sementara objek dengan karakteristik berbeda akan masuk ke cluster lain.

Ada beberapa studi terdahulu telah menggunakan algoritma K-means dalam metode clustering untuk pengelompokan data, salah satunya adalah Mardalius (2018) menggunakannya untuk mengelompokkan data penjualan aksesoris, Mudzakkir (2018) untuk data penjualan produk di PT Advanta Seeds, dan Napitupulu et al. (2020) untuk mengelompokkan dokumen akta kelahiran berdasarkan kecamatan di Kabupaten Simalungun. Tujuan utama dari pengelompokan data ini adalah untuk mendapatkan informasi penting yang dapat mendukung pengambilan keputusan yang lebih baik.

K-Means menawarkan solusi untuk mengidentifikasi pola tersembunyi dalam kumpulan data yang besar dan kompleks. K-Means adalah algoritma pembelajaran tanpa pengawasan (unsupervised learning) yang bekerja dengan mengelompokkan objek data ke dalam k kelompok, di mana setiap objek termasuk dalam kluster dengan pusat (centroid) terdekat. Dengan menerapkan K-Means pada data sekolah menengah di Kabupaten Muna Barat, diharapkan dapat terbentuk kelompok-kelompok sekolah yang homogen dalam karakteristik tertentu.

Penelitian ini bertujuan untuk mengatasi kesenjangan informasi tersebut dengan memanfaatkan algoritma K-Means untuk mengelompokkan data sekolah menengah di Kabupaten Muna Barat. Hasil pengelompokan ini diharapkan dapat memberikan wawasan baru bagi dinas pendidikan setempat dalam merumuskan kebijakan yang lebih terarah, seperti pemerataan kualitas pendidikan, alokasi bantuan, pengembangan program peningkatan mutu guru, atau bahkan perencanaan pembangunan infrastruktur yang lebih efisien berdasarkan karakteristik kluster sekolah.

2. Tinjauan Literatur

2.1. Data Mining

Penambangan data adalah penemuan struktur yang menarik, tak terduga, atau berharga dalam kumpulan data besar (Hand et al., 2001). Data mining atau penambangan data adalah serangkaian langkah untuk menemukan informasi berharga yang sebelumnya tersembunyi dalam basis data.

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan (Mai et al., 2022) yaitu:

- 1) Estimasi mirip dengan klasifikasi, namun variabel targetnya bersifat numerik (bukan kategori). Model dibuat menggunakan data lengkap yang sudah memiliki nilai variabel target sebagai dasar prediksi. Selanjutnya, ketika data baru muncul, nilai variabel target dapat diperkirakan berdasarkan variabel prediktornya.
- 2) Prediksi serupa dengan klasifikasi dan estimasi, namun fokusnya adalah hasil di masa depan. Beberapa algoritma dan teknik dari klasifikasi dan estimasi bisa digunakan untuk prediksi, tergantung pada kasusnya.
- 3) Dalam klasifikasi, variabel tujuannya bersifat kategoris. Contohnya adalah mengelompokkan persediaan menjadi tiga kategori: tinggi, sedang, atau rendah.
- 4) Clustering adalah teknik untuk mengelompokkan data (rekaman, observasi, atau kasus) yang memiliki kemiripan. Sebuah cluster adalah kumpulan rekaman yang saling mirip satu sama lain, namun berbeda dengan rekaman di cluster lain.
- 5) Asosiasi bertujuan untuk mengidentifikasi hubungan antara beberapa kejadian yang terjadi secara bersamaan. Dalam bisnis, ini lebih dikenal sebagai analisis keranjang belanja.

(Silvia Sihombing et al., n.d.) menjelaskan bahwa data mining unggul sebagai alat analisis karena kemampuannya menangani data dalam jumlah besar dan kompleks, serta berbagai tipe atribut. Selain itu, data mining dapat mengolah data secara semi-otomatis, yang berarti sebagian prosesnya otomatis namun tetap membutuhkan masukan parameter manual dari pengguna untuk beberapa teknik. Kelebihan lainnya adalah kemampuannya belajar dari pengalaman atau kesalahan sebelumnya untuk meningkatkan kualitas dan hasil analisis. Namun, menurut (Silvia Sihombing et al., n.d.), kelemahan data mining terletak pada pencarian datanya yang tidak dilakukan secara individual, melainkan dalam bentuk kelompok atau "set individualis" berdasarkan kriteria tertentu.

2.2 Metode Clustering

Clustering adalah salah satu tugas utama dalam data mining yang bertujuan untuk mengelompokkan objek data sedemikian rupa sehingga objek dalam kelompok yang sama memiliki kesamaan yang lebih besar satu sama lain daripada objek dalam kelompok lain. Clustering sering disebut sebagai segmentasi data, di mana data dibagi menjadi segmen-segmen yang homogen.

2.3 Algoritma K-Means

Algoritma K-Means adalah algoritma clustering yang populer dan sederhana yang mengelompokkan data ke dalam k kelompok (clusters) yang telah ditentukan sebelumnya. Prinsip dasarnya adalah mempartisi n observasi ke dalam k klaster, di mana setiap observasi termasuk dalam klaster dengan rata-rata (mean) terdekat. Tahapan umum algoritma K-Means meliputi:

- 1) Inisialisasi Centroid: Memilih secara acak k titik data sebagai pusat klaster awal (centroid).

- 2) Penetapan Klaster: Setiap titik data dialokasikan ke klaster yang centroid-nya paling dekat (sering menggunakan jarak Euclidean).
- 3) Pembaruan Centroid: Centroid setiap klaster dihitung ulang sebagai rata-rata dari semua titik data yang termasuk dalam klaster tersebut. Iterasi: Langkah 2 dan 3 diulang hingga centroid tidak lagi berpindah secara signifikan atau jumlah iterasi maksimum tercapai.

3. Metode

Metode penelitian merupakan cara atau langkah yang digunakan penulis dalam melakukan suatu penelitian melalui tahapan-tahapan yang disusun secara ilmiah sehingga dapat dipergunakan untuk menganalisa, mengerjakan, dan mengatasi masalah yang dihadapi. Pada rancangan penelitian yang digunakan yaitu mempelajari literatur, mengumpulkan data, menganalisa data, menganalisa metode K-Means, implementasi software, menguji hasil, dan menganalisa hasil.

K-Means adalah algoritma clustering yang membagi data menjadi k kelompok berdasarkan kesamaan karakteristik, di mana setiap kelompok memiliki titik pusat (centroid). Tujuan utamanya adalah mengelompokkan objek dengan sifat yang serupa ke dalam satu grup, sedangkan objek dengan sifat berbeda akan masuk ke grup yang lain.

Menurut (Wakhidah, n.d.) langkah-langkah melakukan clustering dengan algoritma K-Means adalah sebagai berikut :

- 1) Menentukan jumlah cluster (k)
- 2) Menentukan Nilai *Centroid*

Dalam menentukan nilai centroid untuk awal iterasi, nilai awal centroid dilakukan secara acak. Sedangkan jika menentukan nilai centroid yang merupakan tahap dari iterasi, maka digunakan rumus sebagai berikut:

$$v_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} \binom{n}{k} x_{kj}$$

Keterangan:

v_{ij} = centroid rata-rata cluster ke-I untuk variabel ke-j

N_i = jumlah data yang menjadi anggota cluster ke-i

i, k = indeks dari cluster

j = indeks dari variabel

X_{kj} = nilai data ke-k yang ada di dalam cluster tersebut untuk variable ke-j

- 3) Menentukan jumlah cluster (k)

Untuk menghitung jarak tersebut dapat menggunakan Euclidean Distance, sebagai berikut:

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

Keterangan:

De = Euclidean Distance

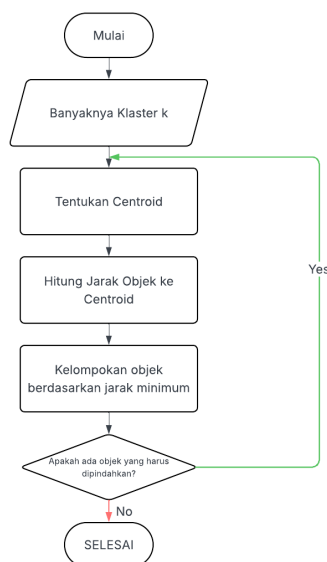
i = banyaknya objek

i = banyaknya objek

(x,y) = koordinat objek

$(s,t) = \text{koordinat centroid}$

- 4) Pengelompokkan objek dengan memperhitungkan jarak minimum objek.
- 5) Kembali ke tahap 2, lakukan perulangan hingga nilai centroid yang dihasilkan tetap dan anggota cluster tidak berpindah ke cluster lain.



Gambar 1. Diagram alir algoritma K-Means

4. Hasil dan Pembahasan

Pada pengumpulan data ada enam parameter yang digunakan dalam pengolahan data yaitu Peserta Didik, Jumlah Guru, Pegawai, Ruang Kelas, Ruangan Laboratorium, dan uangan Perpustakaan.. Data tersebut disajikan dalam bentuk tabel seperti terlihat pada tabel 1 berikut.

Tabel 1. Data Statistik Sekolah

Nama Sekolah	Peserta Didik	Jumlah Guru	Pegawai	R. Kelas	R. Lab	R. Perpus
SMAN 1 Kusambi	448	55	12	15	5	1
SMAN 2 Kusambi	150	33	10	12	3	1
SMKN 1 Muna Barat	303	71	17	21	5	1
SMAN 2 Maginti	184	26	4	6	2	1
SMAN 1 Maginti	265	48	9	27	5	1
SMAS Islam Terpadu Ittibaul Atsar	27	4	2	2	0	0
SMAN 1 Lawa	632	123	21	22	3	1
SMAN 1 Barangka	335	73	12	20	3	1
SMKS Barakati Muna Barat	88	11	5	1	1	1
SMAN 1 Tikep	558	51	7	22	3	1
SMAN 1 Tiworo Tengah	184	33	4	9	4	1
SMKS Kesehatan Muna Barat	14	12	3	3	0	0
SMAN 1 Napano Kusambi	148	27	3	6	4	1
SMKS Maritim	131	30	2	1	4	1

SMKN 2 Muna Barat	213	37	4	14	2	1
SMAN 1 Wadaga	151	49	23	9	5	2
SMKN 3 Muna Barat	139	40	11	17	1	1
SMAN 1 Sawerigadi	114	51	13	12	7	2
SMAN 2 Sawerigadi	85	32	9	9	3	1
SMKS Harapan Muda Barakati	19	4	2	7	0	0
SMKN 4 Muna Barat	138	41	6	9	3	1

Selanjutnya, data ini akan diproses menggunakan algoritma K-Means untuk melakukan clustering. Tujuannya adalah untuk mendapatkan hasil pengelompokan data yang diinginkan. Langkah-langkah clustering dengan algoritma K-Means adalah sebagai berikut:

4.1 . Menentukan Jumlah Cluster (k)

Pengelompokan data pada pengujian ini sebanyak 3 cluster ($k = 3$), diantaranya yaitu cluster 1 menunjukkan sekolah dengan kapasitas besar (C1), cluster 2 menunjukkan sekolah dengan kapasitas kecil (C2), dan cluster 3 menunjukkan sekolah dengan kapasitas sedang (C3).

4.2 . Menentukan Centroid (Pusat Cluster)

Penentuan pusat awal cluster ditentukan secara random sebagai berikut:

Tabel 2. Pusat Cluster Awal

	Peserta Didik	Jumlah guru	Pegawai	R. Kelas	R. Lab	R. Perpus
C1	400	55	12	15	5	1
C2	265	48	9	27	5	1
C3	350	50	11	18	5	1

4.3 . Menghitung jarak setiap objek ke centroid pada masing-masing cluster

Rumus untuk menghitung jarak dari centroid adalah :

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2}$$

Dengan dilakukan perhitungan menggunakan rumus di atas diperoleh tabel hasil perhitungan jarak centroid sebagai berikut:

Tabel 3. Perhitungan jarak centroid iterasi ke-1

Nama Sekolah	Distance From C1	Distance From C2	Distance From C3
SMAN 1 Kusambi	48.0000	183.5511	98.1784
SMAN 2 Kusambi	251.0000	116.9615	200.8233
SMKN 1 Muna Barat	98.6205	45.5302	51.9134
SMAN 2 Maginti	218.2911	86.7179	168.3271
SMAN 1 Maginti	135.7461	0.0000	85.5219
SMAS Islam Terpadu Ittibaul Atsar	376.8594	243.4748	326.8149
SMAN 1 Lawa	242.0372	374.8160	291.5013

SMAN 1 Barangka	67.6609	74.7462	27.6225
SMKS Barakati Muna Barat	315.5012	182.7731	265.5297
SMAN 1 Tikep	158.2972	293.0717	208.0889
SMAN 1 Tiworo Tengah	217.3499	84.4748	167.2603
SMKS Kesehatan Muna Barat	388.7081	254.8235	338.6074
SMAN 1 Napano Kusambi	253.8720	120.8636	203.8185
SMKS Maritim	270.7083	137.8623	220.7533
SMKN 2 Muna Barat	188.0612	55.0273	137.8840
SMAN 1 Wadaga	249.3893	116.2669	199.5695
SMKN 3 Muna Barat	261.4708	126.7281	211.2771
SMAN 1 Sawerigadi	286.0542	151.8420	236.0974
SMAN 2 Sawerigadi	315.9161	181.6150	265.7781
SMKS Harapan Muda Barakati	384.6427	250.8525	334.5220
SMKN 4 Muna Barat	262.5186	128.5107	212.4500

4.4 . Menentukan Cluster

Alokasikan masing-masing objek ke cluster yang paling terdekat berdasarkan nilai minimal dari nilai cluster dan letakkan pada cluster yang sesuai dengan nilai minimal pada iterasi 1.

Tabel 4. Cluster iterasi ke-1

Nama Sekolah	Distance From C1	Distance From C2	Distance From C3	Cluster
SMAN 1 Kusambi	48.0000	183.5511	98.1784	C1
SMAN 2 Kusambi	251.0000	116.9615	200.8233	C2
SMKN 1 Muna Barat	98.6205	45.5302	51.9134	C2
SMAN 2 Maginti	218.2911	86.7179	168.3271	C2
SMAN 1 Maginti	135.7461	0.0000	85.5219	C2
SMAS Islam Terpadu Ittibaul Atsar	376.8594	243.4748	326.8149	C2
SMAN 1 Lawa	242.0372	374.8160	291.5013	C1
SMAN 1 Barangka	67.6609	74.7462	27.6225	C3
SMKS Barakati Muna Barat	315.5012	182.7731	265.5297	C2
SMAN 1 Tikep	158.2972	293.0717	208.0889	C1
SMAN 1 Tiworo Tengah	217.3499	84.4748	167.2603	C2
SMKS Kesehatan Muna Barat	388.7081	254.8235	338.6074	C2
SMAN 1 Napano Kusambi	253.8720	120.8636	203.8185	C2
SMKS Maritim	270.7083	137.8623	220.7533	C2
SMKN 2 Muna Barat	188.0612	55.0273	137.8840	C2
SMAN 1 Wadaga	249.3893	116.2669	199.5695	C2
SMKN 3 Muna Barat	261.4708	126.7281	211.2771	C2
SMAN 1 Sawerigadi	286.0542	151.8420	236.0974	C2

SMAN 2 Sawerigadi	315.9161	181.6150	265.7781	C2
SMKS Harapan Muda Barakati	384.6427	250.8525	334.5220	C2
SMKN 4 Muna Barat	262.5186	128.5107	212.4500	C2

Dari tabel di atas didapatkan data dengan jarak terdekat C1 sebanyak 3, jarak terdekat C2 sebanyak 10, dan jarak terdekat C3 sebanyak 1.

4.2 . Menjumlahkan nilai yang terpilih pada cluster tersebut kemudian membaginya sebanyak jumlah nilai.

Data centroid baru iterasi ke-1 dapat dilihat pada tabel berikut.

Tabel 5. Centroid baru iterasi ke-1

	Peserta Didik	Jumlah guru	Pegawai	R. Kelas	R. Lab	R. Perpus
C1	546	76	13	20	4	1
C2	138	32	7	10	3	1
C3	335	73	12	20	3	1

Dengan menggunakan langkah yang sama seperti sebelumnya untuk menentukan jarak dari centroid dengan menggunakan centroid baru iterasi ke-1, maka berikut hasil perhitungan jarak centroid dan cluster iterasi ke-2.

Tabel 6. Perhitungan jarak centroid dan hasil cluster iterasi ke-2

Nama Sekolah	Distance From C1	Distance From C2	Distance From C3	Cluster
SMAN 1 Kusambi	100.4213	310.5052	114.5513	C1
SMAN 2 Kusambi	398.4521	12.1022	189.4545	C2
SMKN 1 Muna Barat	243.0935	169.7359	32.5269	C3
SMAN 2 Maginti	365.8608	46.3084	158.9686	C2
SMAN 1 Maginti	282.5564	128.7524	74.7462	C3
SMAS Islam Terpadu Ittibaul Atsar	524.4484	115.3761	316.3210	C2
SMAN 1 Lawa	98.1756	502.1863	301.3204	C1
SMAN 1 Barangka	211.0319	201.0731	0.0000	C3
SMKS Barakati Muna Barat	463.0955	55.4998	255.4741	C2
SMAN 1 Tikep	28.8406	420.1852	224.1473	C1
SMAN 1 Tiworo Tengah	364.8599	45.7447	156.8024	C2
SMKS Kesehatan Muna Barat	536.2460	126.3499	327.3255	C2
SMAN 1 Napano Kusambi	401.4118	12.4473	193.2951	C2
SMKS Maritim	418.1492	12.9294	209.5877	C2
SMKN 2 Muna Barat	335.4968	74.9454	127.5970	C2
SMAN 1 Wadaga	396.2097	26.1690	186.2230	C2
SMKN 3 Muna Barat	408.6426	11.3549	198.7939	C2
SMAN 1 Sawerigadi	432.8242	31.6189	222.2769	C2
SMAN 2 Sawerigadi	463.2704	53.4393	253.5961	C2
SMKS Harapan Muda Barakati	532.2241	122.9072	323.8765	C2

SMKN 4 Muna Barat	409.7322	8.8679	199.9750	C2
-------------------	----------	--------	----------	----

Hasil dari cluster iterasi ke-1 dan cluster iterasi ke-2 menunjukkan bahwa SMAN 1 Maginti, SMAN 1 Barangka pindah ke cluster lain.

Maka dilakukan perhitungan kembali sampai data cluster tidak berubah. Selanjutnya mencari centroid dari iterasi ke-2. Maka diperoleh tabel centroid baru iterasi ke-2 sebagai berikut.

Tabel 7. Centroid baru iterasi ke-2

	Peserta Didik	Jumlah guru	Pegawai	R. Kelas	R. Lab	R. Perpustakaan
C1	546	76	13	20	4	1
C2	138	32	7	10	3	1
C3	335	73	12	20	3	1

Dengan menggunakan langkah yang sama seperti sebelumnya untuk menentukan jarak dari centroid dengan menggunakan centroid baru iterasi ke-1, maka berikut hasil perhitungan jarak centroid dan cluster iterasi ke-2.

Tabel 8. Perhitungan jarak centroid dan hasil cluster iterasi ke-3

Nama Sekolah	Distance From C1	Distance From C2	Distance From C3	Cluster
SMAN 1 Kusambi	100.4213	330.1814	147.4777	C1
SMAN 2 Kusambi	398.4521	31.7530	154.5466	C2
SMKN 1 Muna Barat	243.0935	189.5614	8.6603	C3
SMAN 2 Maginti	365.8608	65.1397	124.4642	C2
SMAN 1 Maginti	282.5564	148.5575	39.8079	C3
SMAS Islam Terpadu Ittibaul Atsar	524.4484	95.5831	281.4901	C2
SMAN 1 Lawa	98.1756	521.9896	336.3238	C1
SMAN 1 Barangka	211.0319	220.9031	35.3034	C3
SMKS Barakati Muna Barat	463.0955	36.3995	220.7200	C2
SMAN 1 Tikep	28.8406	439.7973	257.3953	C1
SMAN 1 Tiworo Tengah	364.8599	65.2277	122.1147	C2
SMKS Kesehatan Muna Barat	536.2460	106.5241	292.5286	C2
SMAN 1 Napano Kusambi	401.4118	29.3755	158.5854	C2
SMKS Maritim	418.1492	14.7101	175.0409	C2
SMKN 2 Muna Barat	335.4968	94.6135	92.8906	C3
SMAN 1 Wadaga	396.2097	41.3568	151.7234	C2
SMKN 3 Muna Barat	408.6426	25.1764	163.9085	C2
SMAN 1 Sawerigadi	432.8242	24.5191	187.7765	C2
SMAN 2 Sawerigadi	463.2704	34.2615	218.8196	C2
SMKS Harapan Muda Barakati	532.2241	103.1461	288.9689	C2
SMKN 4 Muna Barat	409.7322	22.6992	165.3209	C2

Dari tabel cluster iterasi ke-2 dan cluster iterasi ke-3 terlihat bahwa masih terdapat sekolah menengah yang berpindah cluster yaitu SMKN 2 Muna Barat C2 pindah ke C3. Maka dari itu dilakukan perhitungan jarak dengan centroid baru.

Tabel 9. Centroid baru iterasi ke-3

	Peserta Didik	Jumlah guru	Pegawai	R. Kelas	R. Lab	R. Perpus
C1	546	76	13	20	4	1
C2	112	28	7	7	3	1
C3	279	57	11	21	4	1

Tabel 10. Perhitungan jarak centroid dan hasil cluster iterasi ke-4

Nama Sekolah	Distance From C1	Distance From C2	Distance From C3	Cluster
SMAN 1 Kusambi	100.4213	336.9257	169.1157	C1
SMAN 2 Kusambi	398.4521	38.4419	131.5375	C2
SMKN 1 Muna Barat	243.0935	196.2344	28.4451	C3
SMAN 2 Maginti	365.8608	71.8197	101.2775	C2
SMAN 1 Maginti	282.5564	155.2884	18.1004	C3
SMAS Islam Terpadu Ittibaul Atsar	524.4484	88.9602	258.3972	C2
SMAN 1 Lawa	98.1756	528.7031	359.2285	C1
SMAN 1 Barangka	211.0319	227.6091	58.1990	C3
SMKS Barakati Muna Barat	463.0955	30.4642	197.5807	C2
SMAN 1 Tikep	28.8406	446.5439	279.0970	C1
SMAN 1 Tiworo Tengah	364.8599	71.9747	98.9324	C2
SMKS Kesehatan Muna Barat	536.2460	99.8030	269.5369	C2
SMAN 1 Napano Kusambi	401.4118	35.9970	135.4349	C2
SMKS Maritim	418.1492	20.5060	151.9840	C2
SMKN 2 Muna Barat	335.4968	101.3716	69.6680	C3
SMAN 1 Wadaga	396.2097	46.9520	129.3952	C2
SMKN 3 Muna Barat	408.6426	31.1161	141.1298	C2
SMAN 1 Sawerigadi	432.8242	24.6416	165.3908	C2
SMAN 2 Sawerigadi	463.2704	27.6960	195.9812	C2
SMKS Harapan Muda Barakati	532.2241	96.5087	265.9044	C2
SMKN 4 Muna Barat	409.7322	28.8456	142.4715	C2

Dari tabel di atas dapat dilihat bahwa posisi cluster pada iterasi ke-4 sama dengan posisi cluster iterasi ke-3, artinya posisi cluster sudah tidak berubah sehingga perhitungan dihentikan.

5. Perbandingan dengan Aplikasi Altair AI Studio

Pada implementasi dan pengujian dilakukan menggunakan aplikasi Altair AI Studio 2025.0.1 dengan tujuan untuk membandingkan hasil pengolahan data secara manual dengan hasil pengolahan data menggunakan aplikasi. Adapun langkah-langkahnya adalah sebagai berikut:

5.1 . Import Data

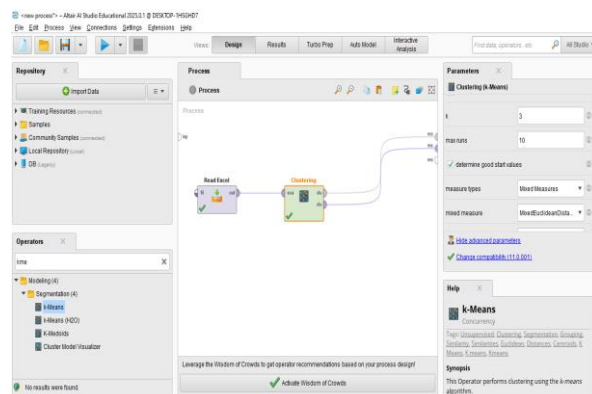
Tahap pertama adalah memasukkan data yang akan diuji dengan klik import data → select the data location → select the cell to import → next → finish.

Row No.	Nama Sekolah	Prestasi Olah Raga	Jumlah Guru	Program	K. Keten	K. Lab	R. Pesisir
1	SMAN 1 Kuasandi	448	55	12	15	5	1
2	SMAN 2 Kuasandi	159	33	10	12	3	1
3	SMAN 1 Muna Barai	303	71	17	21	5	1
4	SMAN 2 Sugeng	164	26	4	6	2	1
5	SMAN 1 Sugeng	265	48	9	27	5	1
6	SMAN Islam Terpadu Muna Barai	27	4	2	2	0	0
7	SMAN 1 Lanes	632	123	21	22	3	1
8	SMAN 1 Barangia	335	73	12	20	3	1
9	SMAN Barakati Muna Barai	88	11	5	1	1	1
10	SMAN 1 Tihap	558	51	7	22	3	1
11	SMAN 1 Tawari Tawari	164	33	4	9	4	1
12	SMAN Khasanah Muna Barai	14	12	3	3	0	0
13	SMAN 1 Nipang Kuasandi	148	27	3	6	4	1
14	SMAN 1 Muna Barai	131	30	2	1	4	1
15	SMAN 2 Muna Barai	213	37	4	14	2	1

Gambar 2. Ini adalah gambar. Skema mengikuti format yang sama.

5.2 . Pengujian data dengan algoritma K-Means

Selanjutnya, data akan diuji menggunakan operator K-Means. Pertama, klik tombol "design" untuk menampilkan layar proses. Kemudian, seret dan lepas data dari Local Repository ke layar proses. Tambahkan operator K-Means dengan mencarinya di kolom pencarian operator, lalu seret dan lepas ke layar proses. Terakhir, hubungkan data ExampleSet dengan operator K-Means agar proses clustering dapat berjalan, seperti yang ditunjukkan pada Gambar 3.



Gambar 3. Tampilan proses K-Means

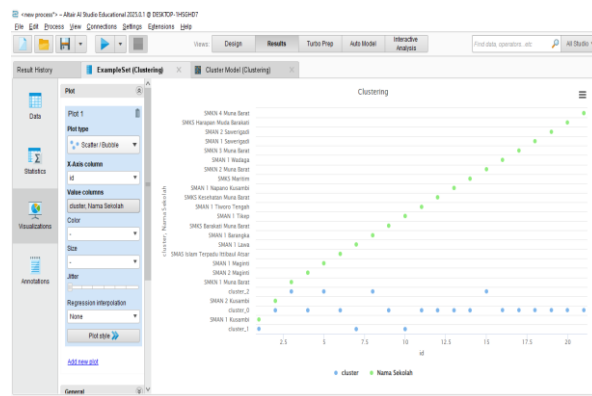
Kemudian, atur algoritma K-Means dengan menentukan nilai k (jumlah cluster) menjadi 3: sekolah besar, kecil dan sedang. Setelah itu, jalankan proses dengan mengklik tombol segitiga di toolbar.

Setelah proses running selesai, maka akan muncul output ExampleSet hasil clustering K-Means, seperti pada gambar 4 berikut:

Row No.	id	cluster	Nama Sekolah	Presentasi	JumlahGuru	Program	R_Kelas	R_Lab	R_Perpus
1	1	cluster_1	SMAN 1 Kusambi	448	55	12	15	5	1
2	2	cluster_3	SMAN 2 Kusambi	150	33	10	12	3	1
3	3	cluster_2	SMAN 1 Muna Barat	303	71	17	21	5	1
4	4	cluster_3	SMAN 2 Mungga	194	26	4	6	2	1
5	5	cluster_2	SMAN 1 Mungga	265	48	9	27	5	1
6	6	cluster_3	SMAN 1 Muna Barat	27	4	2	2	0	0
7	7	cluster_1	SMAN 1 Lanes	632	123	21	32	3	1
8	8	cluster_2	SMAN 1 Baranaga	335	73	12	20	3	1
9	9	cluster_3	SMAN 1 Baranaga Muna Barat	88	11	5	1	1	1
10	10	cluster_1	SMAN 1 Tinas	558	91	7	32	3	1
11	11	cluster_3	SMAN 1 Tinas Tengah	194	33	4	9	4	1
12	12	cluster_2	SMAN 1 Kuantan Muna Barat	14	12	3	3	0	0
13	13	cluster_3	SMAN 1 Kuantan Kusambi	148	27	3	6	4	1
14	14	cluster_3	SMAN 1 Muna Barat	131	36	2	1	4	1
15	15	cluster_2	SMAN 2 Muna Barat	213	37	4	14	2	1

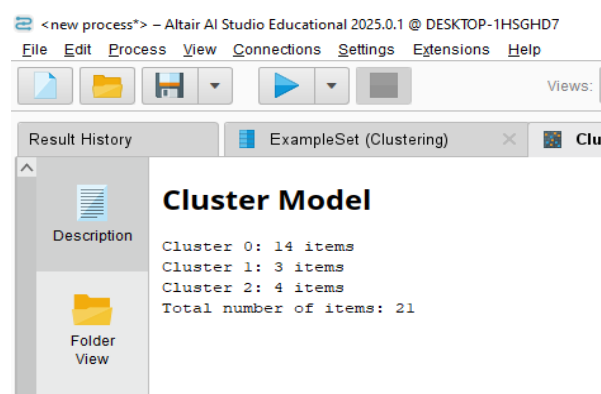
Gambar 4. Tampilan ExampleSet hasil clustering

Tampilan ExampleSet hasil clustering dapat dilihat dalam bentuk scatter plot seperti pada gambar 5 berikut.



Gambar 5. Tampilan scatter plot hasil clustering

Selain itu, hasil clustering juga dapat dilihat pada model cluster yang tampilannya berupa text view yang melihatkan hasil pengelompokan berdasarkan cluster dan jumlah anggotanya, seperti pada gambar 6 berikut:



Gambar 6. Tampilan model cluster

Sehingga dari pengujian data menggunakan bantuan aplikasi Altair AI Studio memiliki perbedaan kelompok cluster dengan perhitungan manual menggunakan Microsoft Excel. Hasil clustering sebagai berikut:

Cluster 0 memiliki 14 data yang berarti merupakan kelompok sekolah kecil, fasilitas minim, butuh intervensi prioritas. Pada perhitungan menggunakan microsoft excel, cluster 0 sebagai C2.

Cluster 1 memiliki 3 data data yang berarti merupakan kelompok sekolah besar dengan fasilitas yang lengkap. Pada perhitungan menggunakan microsoft excel, cluster 1 sebagai C1.

Cluster 2 memiliki 4 data data yang berarti merupakan kelompok sekolah sedang dengan fasilitas yang cukup. Pada perhitungan menggunakan microsoft excel, cluster 2 sebagai C3.

6. Kesimpulan

Berdasarkan hasil pembahasan yang telah diuraikan di atas, dapat ditarik kesimpulan bahwa metode clustering dengan algoritma K-Means dapat digunakan untuk mengelompokkan data sekolah menengah di Kabupaten Muna Barat. Berdasarkan hasil clustering dengan algoritma K-Means, cluster 1 (C1) memiliki jumlah Peserta Didik, Jumlah Guru, Pegawai, R. Kelas, R. Lab, dan R. Perpustakaan yang paling tinggi dibandingkan kluster lainnya. Hal ini mengindikasikan bahwa sekolah-sekolah di Cluster 1 adalah sekolah besar dengan sumber daya yang baik. Kluster 2 (C2) adalah kluster dengan jumlah sekolah terbanyak. Sekolah-sekolah di kluster ini memiliki sumber daya yang beragam, namun secara keseluruhan lebih rendah dibanding Kluster 1 dan 3. Meskipun ada sekolah dengan sumber daya yang memadai, sebagian besar lainnya terbatas pada aspek-aspek tertentu. Cluster 3 (C3) menunjukkan karakteristik sumber daya yang lebih baik dari Cluster 2, mengindikasikan bahwa cluster 3 memiliki sumber daya yang cukup baik dalam beberapa aspek, tetapi tidak sebesar sekolah-sekolah pada Cluster 1.

Kluster ini dapat dimanfaatkan oleh pemerintah atau dinas pendidikan untuk optimalisasi alokasi sumber daya, meliputi guru, fasilitas, dan anggaran. Seperti yang terjadi pada salah satu sekolah dalam Kluster 2 yang tidak memiliki laboratorium atau perpustakaan dapat diprioritaskan dalam program pembangunan fasilitas. Perumusan kebijakan pendidikan dapat disesuaikan dengan kebutuhan spesifik masing-masing kluster. Program pengembangan profesional guru dapat lebih diarahkan pada sekolah di Kluster 2 yang mungkin menghadapi kendala ketersediaan guru. Dengan memahami karakteristik unik setiap kluster, penilaian kinerja sekolah dapat dilakukan secara lebih proporsional dan adil, dengan mempertimbangkan konteks sumber daya yang tersedia.

Referensi

- [1] Fakhri, D. A., Defit, S., & Sumijan. (2021). Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering. *Jurnal Informasi Dan Teknologi*, 160–166. <https://doi.org/10.37034/jidt.v3i3.137>
- [2] Hand, D. J. ., Mannila, Heikki., & Smyth, Padhraic. (2001). *Principles of data mining*. MIT Press.
- [3] Mai, P., Tarigan, S., Tata Hardinata, J., Qurniawan, H., Safii, M., Winanjaya, R., Studi, P., Informasi, S., Tunas, S., & Pematangsiantar, B. (2022). *IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA APRIORI DALAM MENENTUKAN PERSEDIAAN BARANG (STUDI KASUS : TOKO SINAR HARAHAP)* (Vol. 12, Issue 2). <https://jurnal.umj.ac.id/index.php/just-it/index>
- [4] Musharyadi, F., Kom, M., & Padang, S. M. (2017). *TINGKAT PEMAHAMAN MAHASISWA TERHADAP NORMA NORMA AGAMA ISLAM MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING*. 78.
- [5] Mardalius. (2018). Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means. *Jurnal Teknologi dan Sistem Informasi*, 4(2), 123-132
- [6] Napitupulu, F. S., Damanik, I. S., Saragih, I. S., & Wanto, A. (2020). Algoritma K-Means untuk Pengelompokkan Dokumen Akta Kelahiran pada Tiap Kecamatan di Kabupaten Simalungun. *Building of Informatics Technology and Science (BITS)*, 2(1), 55-63.
- [7] Fakhri, D.A., Defit, S., & Sumijan. (2021). Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering. *Jurnal Informasi dan Teknologi*, 3(3), 160-166.
- [8] Fakhri, D.A., Defit, S., & Sumijan. (2021). Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering. *Jurnal Informasi dan Teknologi*, 3(3), 160-166.
- [9] Silvia Sihombing, E., Setiawan Honggowibowo, A., & Nugraheny Jurusan Teknik Informatika Sekolah Tinggi Teknologi Adisutjipto Yogyakarta, D. (n.d.). *Implementasi Data Mining Menggunakan Metode Apriori Pada Transaksi Penjualan Barang (Studi Kasus Di Chorus Minimarket)*.
- [10] Tri, A., Dani, R., Wahyuningsih, S., & Rizki, N. A. (2019). Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu. *Jambura Journal of Mathematics*, 1. <http://ejurnal.ung.ac.id/index.php/jjom,P->
- [11] Wakhidah, N. (n.d.). *CLUSTERING MENGGUNAKAN K-MEANS ALGORITHM (K-MEANS ALGORITHM CLUSTERING)*.
- [12] Yaro, I. (2023). Journal of Social Transformation and Regional Development Relevance of School Facilities and Infrastructure Towards Implementing Education Policies. *JOURNAL OF SOCIAL TRANSFORMATION AND REGIONAL DEVELOPMENT*, 5(2), 61–67. <https://doi.org/10.30880/jstard.2023.05.02.007>
- [13] Kementerian Pendidikan Dasar dan Menengah Republik Indonesia (2025). <https://www.kemendikdasmen.go.id/>