

Analysis Cluster K-Means Terhadap Fasilitas Sekolah Pada Tiap Provinsi Republik Indonesia

Muhammad Faisal^{1*}, Suharmanto² Janu Ilham Saputro³, Wiranti Sri Utami⁴

¹ Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Raharja, Kota Tangerang, Indonesia

² Fakultas Teknologi Informasi, Sistem Informasi, Universitas Graha Ganesha, Kota Tangerang, Indonesia

³ Fakultas Sains dan Teknologi, Sistem Informasi, Universitas Raharja, Kota Tangerang, Indonesia

⁴ Fakultas Teknik, Teknik Informatika, Universitas Cendekia Abditama, Kabupaten Tangerang, Indonesia

Email: ^{1*} muhammad.faisal@raharja.info, ² suharmanto.java@gmail.com, ³ janu@raharja.info, ⁴ wirantisutami@uca.ac.id

(* : coresponding author)

Abstrak- Fasilitas Sekolah merupakan Sarana yang disediakan oleh Pihak Sekolah maupun Perguruan Tinggi untuk menunjang kegiatan dan dapat dimanfaatkan oleh Siswa, Guru, Mahasiswa dan Staff yang berada pada ruang lingkup suatu Pendidikan tertentu. Demi terciptanya Kegiatan Belajar Mengajar (KBM) yang baik dan mendukung proses pengembangan maupun prestasi maka Sekolah atau Perguruan Tinggi yang baik harus memiliki Sarana ruang Kelas, laboratorium, Perpustakaan, Kantin, tempat Ibadah dan Lapangan. Dengan menerapkan *data mining* dan memanfaatkan sumber data yang didapatkan serta penerapan metode *cluster K-Means* dapat ditarik informasi terkait Fasilitas Sekolah. Jumlah *cluster* yang diperoleh adalah 2 *cluster* dengan jumlah kuadrat menurut *cluster* 76,0%.

Kata Kunci: *Data mining*; *Cluster K-Means*; Fasilitas

Abstract- School Facilities are facilities provided by schools or universities to support activities and can be utilized by students, teachers, students and staff within the scope of a particular education. In order to create good teaching and learning activities (KBM) and support the development process and achievements, good schools or universities must have classroom facilities, laboratories, libraries, canteens, places of worship and fields. By applying data mining and utilizing the data sources obtained and the application of the K-Means cluster method, information related to school facilities can be drawn. The number of clusters obtained is 2 clusters with the number of squares according to the cluster of 76.0%.

Keywords: Data mining; K-Means Cluster; Facilities

1. PENDAHULUAN

Pendidikan merupakan kunci utama dalam membuka pintu kemajuan, oleh sebab itu Kegiatan belajar mengajar tidak lepas dari sarana dan prasarana untuk mendukung proses Belajar Mengajar, hal ini termasuk dalam bangunan fisik sebagai ruang kelas, laboratorium, perpustakaan, dan bangunan lainnya sebagai sarana penunjang proses belajar mengajar. Namun pada kenyataannya masih banyak fasilitas Sekolah yang masih jauh dari kata layak terutama didaerah terpencil. Fasilitas yang ada pada sekolah merupakan sarana yang membantu guru, dan siswa dalam memberikan akses materi pembelajaran ataupun informasi terkait pembelajaran yang didapatkan dari Sekolah [1]. Meskipun Fasilitas Sekolah merupakan peran kunci dalam keberhasilan proses belajar mengajar pada kenyataannya masih banyak sekolah yang belum memiliki fasilitas layak. Ketidadaan baik dalam hal pendidikan dapat disebabkan oleh beberapa faktor, misalnya Keterbatasan sarana serta prasarana, keterbatasan Sumber daya alam, dan ketidaksiapan kurikulum untuk menghadapi masa depan, dan tantangan paling besar adalah ketidakmerataan distribusi insfrastuktur pendidikan [2]. Keterbatasan tersebut harus menjadi fokus utama Pemerintah dalam menanggulangi ketimpangan Pendidikan. Fasilitas dalam pengelolaan lembaga pendidikan sekolah sering kali mengalami hambatan hal tersebut dapat mengganggu terlaksananya manajemen sarana dan prasaran fasilitas sekolah, hambatan yang terjadi dapat muncul dari beberapa faktor internal maupun eksternal, kerusakan, dan Minimnya Sumber Daya. [3].

Fasilitas pada dunia pendidikan merupakan sarana yang disediakan oleh Sekolah maupun Perguruan tinggi untuk siswa, guru, mahasiswa, dosen, dan staff yang berada pada lingkungan pendidikan untuk memberikan akses informasi pembelajaran secara bersamaan [4]. Oleh karena itu penulis mencoba melakukan penelitian menggunakan teknik *data mining* dengan menerapkan metode *cluster K-Means* dalam mengelompokkan fasilitas sekolah di tiap Daerah atau Provinsi Republik Indonesia. *Data mining* merupakan proses ekstraksi dari suatu *dataset* yang besar untuk mendapatkan informasi, dan tujuan utaman dari *data mining* adalah melakukan identifikasi pola, hubungan, dan informasi yang mungkin belum ditemui dari *dataset* secara langsung [5]. Analisis *cluster* adalah proses dalam menganalisa kumpulan obyek dalam menemukan kesamaan dan perbedaan, *cluster* bertujuan untuk mengelompokkan dan memahami struktud data, dengan menerapkan metode *cluster K-Means* dapat digunakan untuk menjelaskan atau mendefinisikan algoritma dalam menentukan obyek kedalam *cluster* tertentu berdasarkan nilai terdekat [6].

Algoritma *cluster K-Means* dapat diimplementasikan kedalam penelitian yang berkaitan dengan *dataset* salah satunya adalah analisa telur ayam kampung dan ayam negeri, data komoditas telur ayam tercatat pada Badan Pusat Statistik (BPS) dimana harga telur ayam sering mengalami naik turun dengan memanfaatkan *cluster K-Means* dapat mengetahui Provinsi dengan tingkat Komoditas Telur Ayam Terbesar [7]. Kemudian Penelitian mengenai *Cluster K-Means* untuk mengelompokkan Rumah Tangga dengan status Sewa, dengan menerapkan metode *cluster K-Means* dapat memberikan informasi Provinsi dan Rumah Tangga dengan status Rumah Sewa atau kontrak [8]. Penggunaan metode *cluster K-Means*

pada penelitian ini bertujuan untuk mencari informasi mengenai fasilitas sekolah yang masuk kedalam kelompok layak dan tidak layak berdasarkan *dataset* dari Badan Pusat Statistik (BPS). Berdasarkan hasil refrensi yang penulis buat, penulis dapat memberikan kesimpulan bahwa Analisa yang cocok untuk mengelompokkan Fasilitas sekolah Provinsi di setiap Republik Indonesia adalah dengan menerapkan Metode K-Means.

2. METODOLOGI PENELITIAN

2.1 Dataset Penelitian

Untuk mendukung proses penelitian ini penulis menggunakan *dataset* yang berasal dari Badan Pusat Statistik (BPS), data ini akan diolah menggunakan teknik *data mining* dengan menerapkan Metode *cluster* K-Means. Hasil data ini bertujuan untuk menggali informasi mengenai fasilitas Sekolah pada tiap Provinsi. Data yang akan digunakan dapat dilihat pada tabel dibawah ini.

Tabel 1. Data Fasilitas Sekolah & Perguruan Tinggi

Provinsi	SD	SMP	SMU	SMK	Perguruan Tinggi
Aceh	3382	1421	735	205	119
Sumatera Utara	5003	2319	1147	712	202
Sumatera Barat	1256	808	415	180	102
Riau	1811	1210	631	259	80
Jambi	1484	810	393	169	42
Sumatera Selatan	2996	1410	720	252	99
Bengkulu	1187	488	194	93	27
Lampung	2537	1437	695	405	96
Kep. Bangka Belitung	391	217	85	49	16
Kep. Riau	408	249	132	62	33
DKI Jakarta	264	255	219	207	129
Jawa Barat	5948	4314	2191	2070	477
Jawa Tengah	8423	3692	1343	1204	314
DI Yogyakarta	437	310	148	139	66
...	...	...	...	...	...
Papua Pegunungan	530	169	48	15	8

Data yang sudah didapatkan tersebut kemudian akan diproses menggunakan *tools* atau *Software* Rstudio. Sebelum dilakukan analisa dengan *Software* Rstudio data akan di *upload* ke Rstudio dan kemudian akan diolah.

2.2 Data Mining

Data mining merupakan bagian dari *Knowledge discovery*, *Knowledge extraction*, *data/pattern analysis*, *information harversting*, dan lain-lain. Pengertian lain dari teknik *data mining* merupakan suatu proses untuk menganalisa pola data yang tersembunyi dan menurut sebagian orang dapat diolah menjadi informasi yang berguna [9]. Dalam *data mining* terdapat istilah atau teknik *data mining* yaitu *Cluster analysis*, *cluster analysis* merupakan teknik untuk mengetahui kelompok atau obyek yang memiliki kesamaan satu dengan lainnya dalam kelompok namun memiliki perbedaan dengan obyek atau kelompok lainnya [10].

2.3 Cluster K-Means

Cluster K-Means merupakan merupakan pengelompokkan algoritma yang iteratif penerapannya digunakan sebagai partisi *dataset* kedalam sejumlah *cluster* data yang telah ditetapkan [11]. Algoritma ini telah banyak digunakan

untuk penelitian yang berkaitan dengan *dataset* salah satu contohnya adalah dalam Bidang Migrasi Penduduk, dengan menerapkan *cluster* K-Means untuk pengelompokkan data migrasi dengan cara mengklasifikasikan tiap kecamatan terhadap *cluster* yang telah ditetapkan [12]. Kemudian Penelitian *cluster* K-Means untuk Prediksi jenis penyakit di Provinsi Jawa Timur, dengan menerapkan *cluster* K-Means bertujuan untuk mengembangkan sistem informasi dan mengimplementasikan pengelompokkan dan prediksi jumlah penyakit tiap Kota maupun Kabupaten pada Provinsi Jawa Timur [13].

Kemudian mendukung penelitian ini penulis menggunakan metode *elbow*, metode *elbow* adalah sebuah metode yang digunakan sebagai penentuan jumlah optimal dari *cluster* yang akan dibentuk dengan menerapkan metode *cluster* K-Means [14]. Metode *Elbow* pada penelitian ini digunakan untuk memilih nilai k terkecil dari data yang didapatkan yang masih memiliki nilai *within* atau terendah dan nilai perbandingan *sum of square errors* (SSE) dapat ditarik kesimpulan dari total *cluster* yang membentuk grafik siku pada suatu pola titik [15]. Persamaan atau rumus dari metode *elbow* dapat dilihat pada dibawah ini.

$$SSE = \sum_k^k = 1 \sum x_i = S_k || N_i - C_k ||$$

Selain menggunakan metode *Elbow* penulis juga menggunakan metode Korelasi *Pearson*, Korelasi *pearson* digunakan untuk menganalisa hubungan antara variabel masukan dan variabel keluaran [16]. Dalam penelitian ini variabel masukan yang akan digunakan adalah *dataset* fasilitas sekolah yang ada di Republik Indonesia. Korelasi *pearson* dapat dihitung menggunakan Rumus berikut ini.

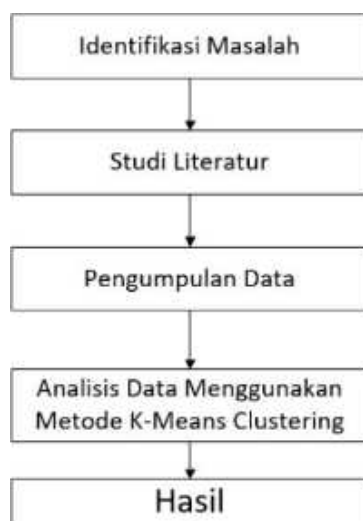
$$r_{xy} = \frac{\sum xy}{(n-1)S_x S_y} \dots (1)$$

Setelah menggunakan Pendekatan *Elbow* Penulis juga menggunakan pendekatan *Silhouette coefficient*. Nilai *Silhouette coefficient* terletak pada nilai -1 hingga 1, maksudnya adalah jika mendekati nilai 1 maka semakin baik pengelompokkan dalam suatu *cluster* sebaliknya jika mendekati nilai -1 data semakin buruk kualitasnya dalam suatu *cluster*. Rumus tersebut dapat dilihat dibawah ini.

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

2.4 Tahapan Penelitian

Tahap selanjutnya Adalah merumuskan Langkah-langkah yang akan dilakukan dalam penelitian ini, untuk menjelaskan alur penelitian secara lengkap berikut ini merupakan Gambaran penelitian menggunakan metode *Cluster* K-Means.



Gambar 1. Langkah-langkah Penelitian

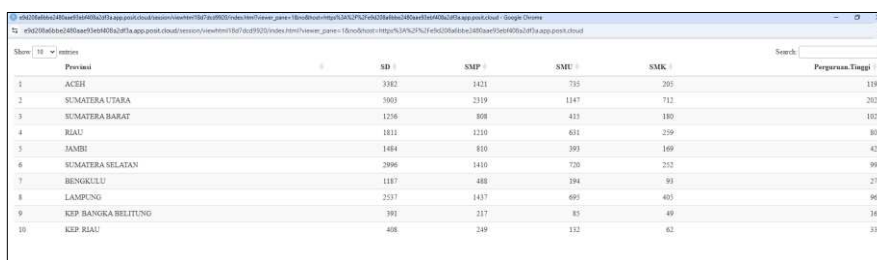
Langkah berikutnya penulis akan melakukan Analisa menggunakan perhitungan statistik, untuk membantuk proses perhitungan statistik penulis menggunakan *software* *Posit Cloud*. *Posit Cloud* merupakan *software* yang menggunakan Bahasa pemrograman R, *software* ini bersifat *open source* dapat digunakan asalkan terkoneksi dengan jaringan internet.

3. HASIL DAN PEMBAHASAN

Setelah mendapatkan *dataset* yang diperoleh dari Badan Pusat Statistik (BPS) dalam bentuk *file* Excel, data di *import* kedalam *software* Posit Cloud.

3.1 Import Data

Data yang akan di *import* menggunakan fungsi *read_excel()*, fungsi tersebut merupakan bagian yang ada dalam *software* Posit Cloud. Hal ini dilakukan karena *file* yang didapatkan oleh penulis adalah *file* dalam bentuk Excel. Proses tersebut dapat dilihat pada gambar dibawah ini.



	Provinsi	SD	SMP	SMU	SMK	Perguruan Tinggi
1	ACEH	3382	1421	735	205	119
2	SUMATERA UTARA	3063	2319	1147	712	202
3	SUMATERA BARAT	1256	808	413	180	102
4	RIAU	1811	1210	611	239	80
5	JAMBE	1484	810	393	149	42
6	SUMATERA SELATAN	2966	1410	720	212	99
7	BENGKULU	1187	488	294	93	27
8	LAMPUNG	2537	1437	693	405	96
9	KEP. BANGKA BELITUNG	391	217	85	49	16
10	KEP. RIAU	408	249	112	62	33

Gambar 2. Data Fasilitas Sekolah yang telah di *import* ke *software* Posit Cloud

3.2 Analisa Data

Setelah data di *import* langkah selanjutnya adalah melakukan proses statistik untuk mendapatkan ringkasan informasi dari data yang digunakan, fungsi yang digunakan adalah *summary*, Proses tersebut dapat dilihat pada gambar dibawah ini.

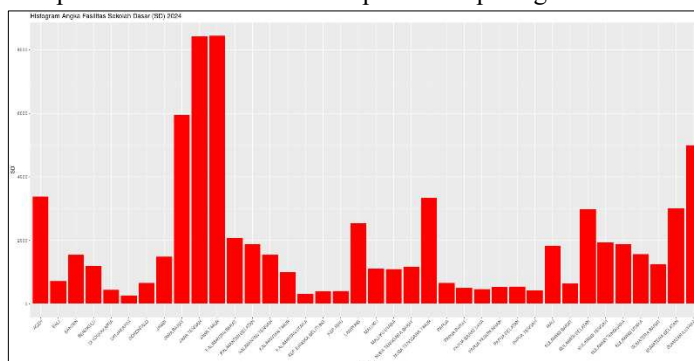
```
> ##statistik deskriptif
> statdesS <- summary(dataLatihan)
> statdesS
```

Provinsi	SD	SMP	SMU	SMK	Perguruan Tinggi
Length:38	Min. : 264.0	Min. : 110.0	Min. : 32.0	Min. : 15.00	Min. : 8.00
Class :character	1st Qu.: 566.5	1st Qu.: 268.8	1st Qu.: 136.0	1st Qu.: 69.75	1st Qu.: 23.50
Mode :character	Median :1221.5	Median : 746.5	Median : 284.0	Median : 152.50	Median : 46.50
	Mean :1907.1	Mean :1022.0	Mean : 477.1	Mean : 289.74	Mean : 86.55
	3rd Qu.:2041.2	3rd Qu.:1246.8	3rd Qu.: 658.0	3rd Qu.: 257.25	3rd Qu.: 98.25
	Max. :8442.0	Max. :4905.0	Max. :2587.0	Max. :2070.00	Max. :477.00

Gambar 3. Proses Statistik menggunakan fungsi *summary*

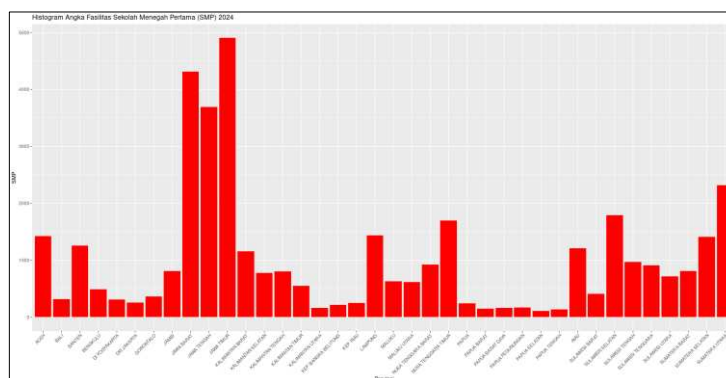
Dari informasi diatas dapat dilihat terdapat 38 *character* dari 5 variabel, dimana variabel SD nilai Median-nya adalah 1221,5, nilai Mean-nya adalah 1907,1, dan nilai Max-nya adalah 8442,0. Kemudian untuk SMP nilai Median-nya adalah 746,5, nilai Mean-nya adalah 1022,0, dan nilai Max-nya adalah 4905,0. Selanjutnya untuk nilai SMU nilai Median-nya adalah 284,0, nilai Mean-nya adalah 477,1, dan nilai Max-nya adalah 2587,0. Selanjutnya untuk nilai SMK nilai Median-nya adalah 152,50, nilai Mean-nya adalah 289,74, dan nilai Max-nya adalah 2070,0. Selanjutnya untuk nilai Perguruan Tinggi nilai Median-nya adalah 46,50, nilai Mean-nya adalah 86,55, dan nilai Max-nya adalah 477,00.

Setelah membuat proses statistik penulis mencoba melakukan analisa fasilitas sekolah dalam bentuk visualisasi Histogram untuk Sekolah Dasar (SD). Tampilan visualisasi tersebut dapat dilihat pada gambar dibawah ini.



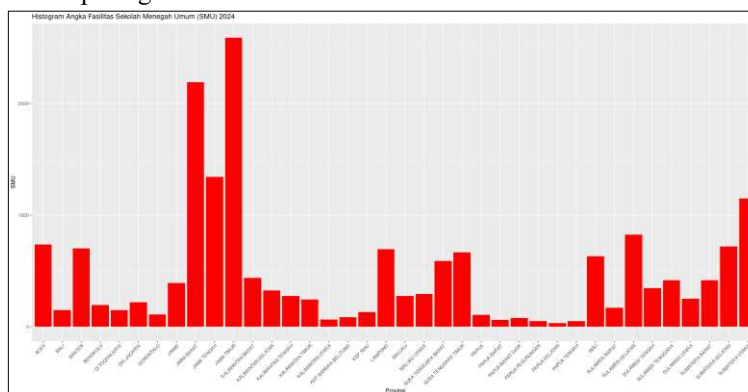
Gambar 4. Visualisasi Histogram Fasilitas Sekolah Dasar Tahun 2024

Dari gambar visualisasi histogram diatas dapat diperhatikan Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur memiliki tingkat Fasilitas yang tinggi. Oleh sebab itu Provinsi yang grafiknya rendah wajib menjadi perhatian khusus untuk meningkatkan Fasilitas Sekolah Dasar. Kemudian selanjutnya penulis akan membuat tampilan visualisasi untuk Sekolah Menengah Pertama (SMP) Tampilan tersebut dapat dilihat pada gambar dibawah ini.



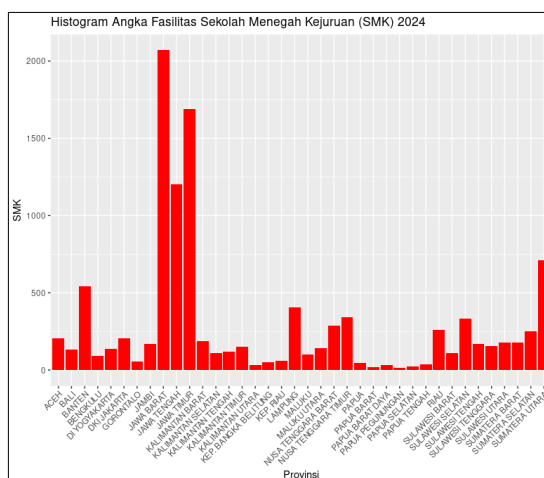
Gambar 5. Visualisasi Histogram Fasilitas Sekolah Menengah Pertama Tahun 2024

Dari tampilan visualisasi Sekolah Menengah Pertama (SMP) dapat dilihat data histogram yang tinggi berada pada Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur memiliki tingkat Fasilitas yang tinggi. Provinsi lainnya yang rendah menjadi perhatian khusus pemerintah dalam memberikan fasilitas yang lebih layak. Setelah membuat tampilan visualisasi Sekolah Menengah Pertama (SMP) kemudian penulis membuat tampilan visualisasi Sekolah Menengah Umum (SMU) Tampilan tersebut dapat dilihat pada gambar dibawah ini.



Gambar 6. Visualisasi Histogram Fasilitas Sekolah Menengah Umum Tahun 2024

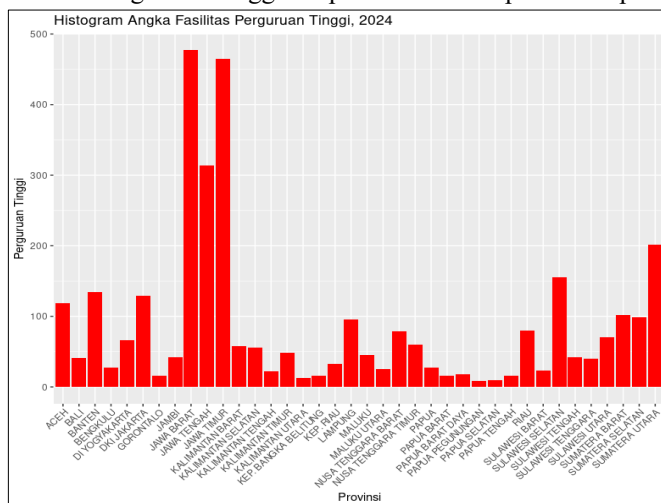
Dari tampilan visualisasi Sekolah Menengah Umum (SMU) dapat dilihat data histogram yang tinggi berada pada Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur memiliki tingkat Fasilitas yang tinggi. Provinsi lainnya yang rendah menjadi perhatian khusus pemerintah dalam memberikan fasilitas yang lebih layak. Selanjutnya penulis akan membuat tampilan visualisasi histogram dari data Sekolah Menengah Kejuruan (SMK) tampilan tersebut dapat dilihat pada gambar dibawah ini.



Gambar 7. Visualisasi Histogram Fasilitas Sekolah Menengah Kejuruan Tahun 2024

Dari tampilan visualisasi Sekolah Menengah Kejuruan (SMK) dapat dilihat data histogram yang tinggi tetap berada pada Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur memiliki tingkat Fasilitas yang tinggi. Provinsi lainnya yang rendah

menjadi perhatian khusus pemerintah dalam memberikan fasilitas yang lebih layak. Kemudian penulis akan membuat tampilan visualisasi histogram untuk Perguruan Tinggi tampilan tersebut dapat dilihat pada gambar dibawah ini.



Gambar 8. Visualisasi Histogram Fasilitas Perguruan Tinggi Tahun 2024

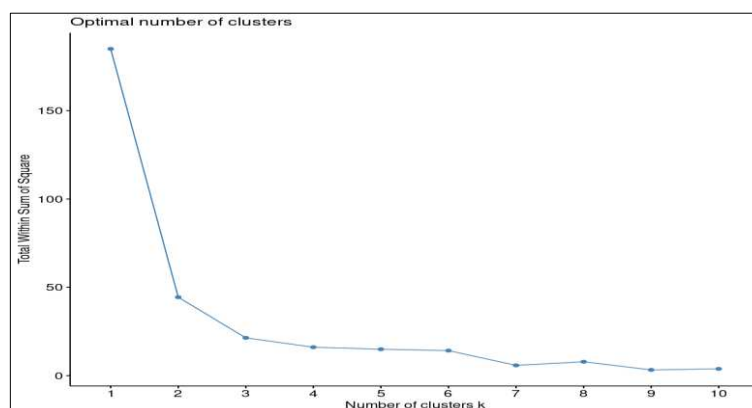
Dari tampilan visualisasi histogram diatas Provinsi yang memiliki fasilitas lebih baik tetap berada pada Jawa Barat, Jawa Tengah, dan Jawa Timur. Masih ada beberapa Provinsi yang kategorinya rendah, Provinsi yang rendah tersebut menjadi perhatian khusus pemerintah dalam memberikan fasilitas yang lebih layak. Kemudian penulis akan melakukan uji Korelasi terhadap Fasilitas Sekolah dari tingkat SD sampai dengan Perguruan Tinggi berdasarkan data yang diperoleh menggunakan metode Pearson pada pemrograman R. Proses tersebut dapat dilihat pada gambar dibawah ini.

```
> kmoS <- KMO(dataLatihan[,2:6])
> korelasi <- cor(dataLatihan[,2:6], method = 'pearson')
> korelasi
```

	SD	SMP	SMU	SMK	Perguruan Tinggi
SD	1.0000000	0.9666192	0.9197459	0.8697731	0.8823789
SMP	0.9666192	1.0000000	0.9832400	0.9496437	0.9493986
SMU	0.9197459	0.9832400	1.0000000	0.9531978	0.9613564
SMK	0.8697731	0.9496437	0.9531978	1.0000000	0.9755115
Perguruan Tinggi	0.8823789	0.9493986	0.9613564	0.9755115	1.0000000

Gambar 9. Proses Uji Korelasi menggunakan Metode Pearson

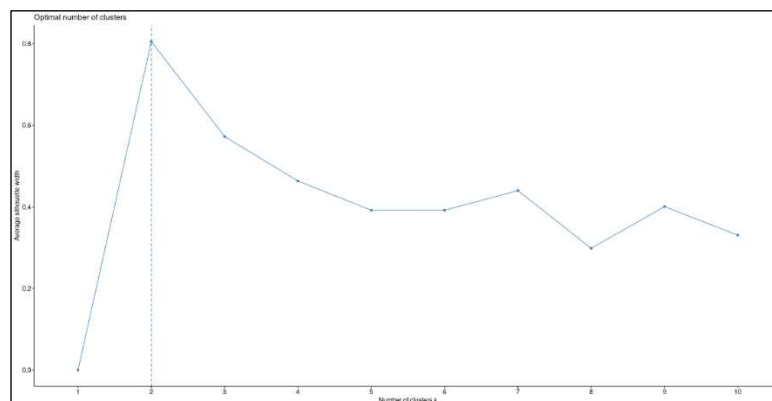
Proses selanjutnya penulis akan menentukan banyaknya *cluster* yang akan digunakan pada penelitian ini. Proses tersebut menggunakan metode *elbow*. Proses dalam menentukan jumlah *cluster* yang akan dipakai pada penelitian ini adalah dari titik siku yang mengalami penurunan drastis diikuti dengan penurunan yang tidak terlalu drastis selanjutnya. Metode *elbow* yang digunakan pada penelitian ini dengan menerapkan fungsi WSS (*Within-Cluster Sum of Square*). Gambar titik siku dari metode *elbow* dapat dilihat pada gambar dibawah ini.



Gambar 10. Cluster Optimal menggunakan metode *elbow*

Berdasarkan hasil analisa menggunakan metode *elbow* dapat diamati penurunan yang terjadi berada pada titik 2 namun terjadi penurunan drastis kembali berada pada titik 7. Penulis menyimpulkan *cluster* yang akan digunakan pada penelitian ini adalah 2 *cluster*.

Selain menggunakan metode *elbow* Penulis juga Menggunakan Metode *Silhouette coefficient* sebagai validitas analisis *cluster* fungsi yang digunakan pada *Software Rstudio* adalah “*silhouette*”. Tampilan tersebut dapat dilihat pada gambar dibawah ini.

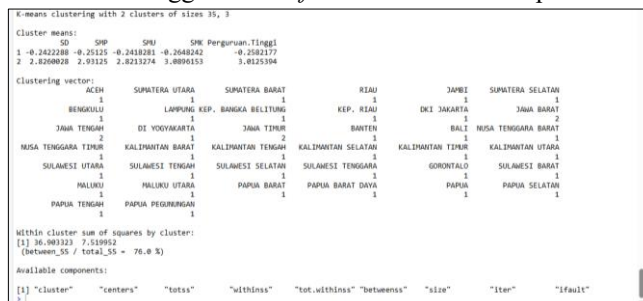


Gambar 11. Hasil Validitas *cluster* menggunakan metode *Silhouette coefficient*

Berdasarkan hasil validitas *cluster* menggunakan metode *Silhouette coefficient*, *cluster* optimal digunakan pada penelitian ini adalah sebanyak 2 *cluster*. Sehingga penulis membuat kesimpulan akhir bahwa jumlah *cluster* yang digunakan pada penelitian ini adalah 2 *cluster*.

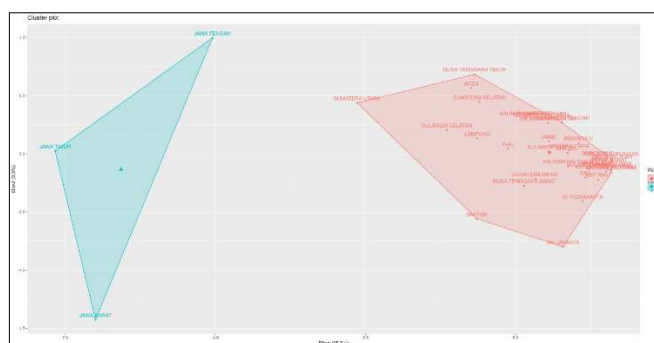
3.3 Analisa Cluster

Analisa *cluster* yang digunakan pada penelitian ini adalah menggunakan metode *cluster K-Means*, dalam penerapan Algoritma *cluster K-Means* hal yang perlu diperhatikan adalah penentuan nilai *centroid* awal menggunakan metode *K-Means*. Pertama-tama dimulai dengan menentukan *centroid* awal dengan cara *random* atau acak terhadap data yang akan digunakan. Proses penggunaan *centroid* awal secara acak dapat menyebabkan nilai yang dihasilkan dapat berbeda-beda jika hanya melakukan proses *centroid* hanya sekali. Oleh sebab itu perlu dilakukan pengulangan dengan menggunakan Algoritma *K-Means* beberapa kali dengan menempatkan nilai awal yang berlainan untuk mendapatkan nilai *cluster* yang optimal. Penerapan Algoritma *K-Means* menggunakan *software Posit Cloud* dapat dilihat pada gambar dibawah ini.



Gambar 12. Hasil analisa menggunakan Algoritma *K-Means*

Selanjutnya penulis akan membuat tampilan visualisasi dalam bentuk *cluster Plot* yang akan memberikan informasi Provinsi *cluster 1* dan *cluster 2* berdasarkan data yang diperoleh dari fasilitas Sekolah. *Cluster plot* tersebut dapat dilihat pada gambar dibawah ini.



Gambar 13. *Cluster Plot* Fasilitas Sekolah pada tiap Provinsi

Berdasarkan *cluster* plot diatas dapat dilihat Provinsi dengan Fasilitas Sekolah terbaik berada pada *cluster* 1 yang beradap pada Provinsi Jawa Timur, Kawa Tengah, dan Jawa Barat. Sedangkan sisanya berada pada *cluster* 2, untuk dapat melihat lebih *detail* Provinsi yang berada pada *cluster* 1 dan *cluster* 2 dapat dilihat pada tabel dibawah ini.

Cluster	Mean_SD	Mean_SMP	Mean_SMT	Mean_SMK	Mean_PengajaranTinggi	Total_Mean	Provinsi
1	1	1028	761	341	275	87	2762
2	2	989	4384	2040	3033	87	13489

Tabel 1. *Cluster* Fasilitas Sekolah pada tiap Provinsi

4. KESIMPULAN

Dengan dilakukannya penelitian ini diharapkan dapat menjadi informasi terhadap data fasilitas Sekolah yang perlu diperhatikan agar kegiatan proses belajar mengajar menjadi lebih efektif, dan dapat dijadikan sebagai bahan evaluasi oleh pihak terkait mengenai sarana prasarana fasilitas Sekolah. Berdasarkan Penelitian yang sudah dibahas diatas Provinsi yang fasilitasnya masuk kedalam kategori tertinggi berada pada Provinsi Jawa Barat, Jawa Tengah, dan Jawa Timur.

REFERENCES

- [1] A. Dila, F. N. Baihaqi, S. Habibah and A. Marini, "Studi Pustaka tentang Peran Fasilitas Sekolah dalam Mendukung Efektivitas Pembelajaran di Sekolah Dasar," *Jurnal Pendidikan Guru Sekolah Dasar*, vol. 3, no. 1, pp. 1-9, 5 Desember 2024.
- [2] M. S. Ruuhulhaq, B. Waluya and S. Himayah, "Analisis Ketersediaan dan Keterjangkauan Fasilitas Pendidikan SMP Negeri Menggunakan Sistem Informasi Geografis di Kota Cimahi," *Jurnal Pendidikan Geografi Undiksha*, vol. 12, no. 2, pp. 227-238, 19 Agustus 2024.
- [3] A. Nikita, N. P. Lubis and S. Fauziah, "Upaya Manajemen Sekolah dalam Menghadapi Hambatan Sarana Prasarana Pendidikan," *Jurnal Bintang Pendidikan dan Bahasa*, vol. 1, no. 3, pp. 1-9, 30 Oktober 2023.
- [4] S. H. Daulay, S. F. Fitriani and E. W. Ningsih, "Pengaruh Fasilitas Sekolah terhadap Kemampuan dan Motivasi Belajar Siswa," *EDUKATIF: JURNAL ILMU PENDIDIKAN*, vol. 4, no. 3, pp. 3731-3738, 2022.
- [5] P. W. Rahayu, "Buju Ajar Data Mining," in *Data Mining and Knowledge Discovery Process*, Bandung, PT. Sonpedia Publishing Indonesia, 2024, pp. 1-11.
- [6] M. Wahyudi, M. R. Saragih and S. , *Data Mining: Perepana Algoritma k-Means Clustering dan K-Medoids Clustering*, Jakarta: Yayasan Kita Menulis, 2020.
- [7] M. Faisal, S. W. S. Utami and N. Pratiwi, "Pemanfaatan Data Mining Menggunakan Metode K-Means Untuk Analisa Komoditas Telur Ayam," *Bulletin of Information Technology (BIT)*, vol. 5, no. 4, pp. 247-254, 2024.
- [8] M. Faisal, W. S. Utami and J. I. Saputro, "PENERAPAN DATA MINING MENGGUNAKAN METODE K-MEANSUNTUK CLUSTERRUMAH TANGGA DENGAN STATUS KEPEMILIKAN RUMAH SEWA," *ICIT Journal*, vol. 11, no. 2, pp. 227-237, 2025.
- [9] M. Arhami and M. Nasir, *Data Mining Algoritma dan Implementasi*, Yogyakarta: Penerbit Andi, 2020.
- [10] M. "Data Mining dan Aplikasinya," in *Pengantar Data Mining*, Bandung, CV. Widina Media Utama, 2021, pp. 2-10.
- [11] M. Ulfah and A. S. Irtwaty, "Penerapan Data Mining Clustering Menggunakan Metode K-Means Dalam Pengelompokan Buku Perpustakaan Politeknik Negeri Balikpapan," *Fidelity : Jurnal Teknik Elektro*, vol. 4, no. 3, pp. 62-68, 2022.
- [12] N. N. Afidah and M. , "Penerapan Metode Clustering dengan Algoritma K-Means untuk Pengelompokan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang," in *PRISMA, Prosiding Seminar Nasional Matematika*, Semarang, 2023.



- [13] S. A. D. Darmawan and K. , "PENERAPAN METODE K-MEANSCLUSTERINGDAN SIMPLE MOVING AVERAGEUNTUK MEMPREDIKSI JENIS PENYAKIT DI PROVINSI JAWA," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 1, no. 4, pp. 877-886, 2024.
- [14] N. A. Maori, "METODE ELBOW DALAM OPTIMASI JUMLAH CLUSTER PADA K-MEANS CLUSTERING," vol. 14, no. 2, pp. 277-287, Jurnal SIMETRIS.
- [15] M. Faisal, S. W. S. Utami and N. Pratiwi, "Pemanfaatan Data Mining Menggunakan Metode K-Means Untuk Analisa Komoditas Telur Ayam," *Bulletin of Information Technology (BIT)*, vol. 5, no. 4, pp. 247-254, 2024.
- [16] D. Leni, A. Dwiharzandis, R. Sumiati, H. and S. Afriyani, "Seleksi Fitur Berdasarkan Korelasi Pearson dalam Pemodelan Efisiensi Energi Bangunan," *Jurnal Teknik Sains*, vol. 8, no. 2, pp. 104-115, 2023.

