

## Pengelompokan Kebersihan Stasiun Wilayah Daop 4 Semarang Berdasarkan Nilai SLA Menggunakan Analisis Cluster dengan Metode K-Means

Adiyah Mahiruna<sup>1✉</sup>, Heni Sulistyowati<sup>2</sup>, Atika Nurani Ambarwati<sup>3</sup>

<sup>1,2</sup>Institut Teknologi Statistika dan Bisnis Muhammadiyah Semarang, Indonesia

<sup>3</sup>Institut Teknologi Statistika dan Bisnis Muhammadiyah Semarang, Indonesia

✉Corresponding Author: adiyah.mahiruna@itesa.ac.id

### ABSTRAK

Kereta api merupakan salah satu moda transportasi nasional yang mempunyai karakteristik kapasitas angkutan massal, mempunyai keunggulan tersendiri, dan terintegrasi dengan moda transportasi lainnya. Kereta api dalam hal penggunaan teknologi sudah sangat mengikuti, walaupun teknologi sangat penuh resiko dalam hal integritas [1]. Stasiun merupakan tempat operasional kereta api, oleh karena itu stasiun yang bersih dan nyaman akan memberikan kenyamanan bagi masyarakat. memberikan kesan positif kepada penumpang, dengan Penelitian ini bertujuan untuk mengetahui pola pengelompokan stasiun di kawasan Daop 4 Semarang berdasarkan tingkat kebersihannya. Sehingga tidak ada kesenjangan kebersihan di setiap stasiun di kawasan Daop 4 Semarang. Metode yang digunakan dalam penelitian ini adalah algoritma K-Means, yaitu metode pengelompokan data yang membagi data menjadi beberapa cluster secara iteratif berdasarkan karakteristik yang serupa. Data yang digunakan dalam penelitian adalah data SLA (Service Level Agreement) kebersihan yang diambil setiap hari kemudian dijumlahkan dan kemudian digunakan untuk clustering stasiun di wilayah Daop 4 Semarang. Hasil penelitian menunjukkan jumlah cluster sebanyak 4 cluster dengan karakteristik pada cluster 1 lebih tinggi pada ketiga variabel pada cluster 2 dan 3 dengan masing-masing cluster berada pada rata-rata, dan pada cluster 4 tidak mempunyai nilai yang rendah untuk cluster tersebut.

**Kata Kunci:** cluster, k-means, train station

### A. Pendahuluan

Kebersihan lingkungan mempunyai arti sebuah keadaan bebas dari kotoran, termasuk di antaranya, debu, sampah, dan bau [2] sedangkan menurut peraturan Daerah Nomor 4 tahun 2000 tentang penyelenggaraan kebersihan adalah “terciptanya suatu keadaan lingkungan yang bersih, rapih, indah dan nyaman untuk menciptakan kesejahteraan masyarakat dalam lingkungan yang sehat”. Oleh karena itu memperhatikan kebersihan lingkungan berdampak baik pada banyak aspek selain dalam kesehatan juga berdampak pada kenyamanan bersama.

Dalam penelitian ini kebersihan yang dimaksud adalah kebersihan yang ada di stasiun, Menurut Peraturan Menteri 48 tahun 2015 kereta api adalah sarana perkeretaapian dengan tenaga gerak sedangkan stasiun dan fasilitas operasi kereta api adalah prasarananya [3]. Stasiun Kereta Api merupakan prasarana kereta api sebagai tempat pemberangkatan dan pemberhentian kereta api. Peraturan Menteri No. 33 TAHUN 2011 Pasal (3) Stasiun kereta api sebagaimana dimaksud dalam Pasal 1 menurut jenisnya terdiri atas: (a) stasiun penumpang; (b) stasiun barang; dan/atau (c) stasiun operasi. Gedung stasiun kereta api merupakan bagian dari stasiun kereta api yang digunakan untuk melayani pengaturan perjalanan kereta api dan pengguna jasa kereta api [4].

Pada mulanya sejarah perkeretapiian dimulai Pada tahun 1864, kereta api pertama di Indonesia, yang menghubungkan Semarang dengan Tanggung, dibangun oleh Nederlandsch-Indische Spoorweg Maatschappij (NIS). Kemudian, pengembangan jalur kereta api terus berlanjut dengan pembangunan jalur-jalur baru yang menghubungkan berbagai kota di Jawa dan Sumatera. Selain itu menurut badan pusat statistika adanya penurunan penumpang sebanyak 4,93% yang menggunkan tranportasi kereta api, oleh karena itu perlu adanya peningkatan pada kualitas dari PT. kereta api Indonesia terlebih Di wilayah daerah Operasional (Daop) 4 Semarang yang dimana letak dari Daop 4 Semarang yang strategis karena berada di Jawa Tengah yang termasuk memiliki penduduk terpadat di Indonesia. Daop 4 Semarang memiliki 43 stasiun, secara umum jalur keberangkatan kereta mencakup sebagian besar wilayah Jawa Tengah bagian utara. Dari 43 stasiun yang ada di wilayah Daop 4 memiliki lebih banyak stasiun kecil, peneliiian ini dilakukan untuk mengetahui bagaimana pengelompokan yang ada agar tidak adanya kesenjangan yang dalam kebersihan setiap stasiun baik stasiun besar maupun kecil. Adapun kota besar yang dilalui oleh jalur kereta api di wilayah ini antara lain Semarang, Pekalongan, Tegal dan Purwokerto sedangkan jalur kereta api membentang dari arah barat hingga ke arah timur.

Penelitian ini berfokus pada pengelompokan kebersihan stasiun di wilayah Daop 4 Semarang berdasarkan nilai Service Level Agreement (SLA) kebersihan. Tujuannya adalah untuk mengidentifikasi tingkat kebersihan setiap stasiun dan mengetahui karakteristik dari setiap cluster stasiun sehingga upaya peningkatan kebersihan dapat dilakukan lebih efisien. Nilai SLA kebersihan itu sendiri adalah nilai yang diambil dari kegiatan pembersihan di stasiun yang dilakukan setiap hari, adapun beberapa tempat yang dinilai atau ditinjau dalam SLA kebersihan ini. Dengan adanya nilai SLA kebersihan ini akan dilakukan pengelompokan atau clustering stasiun sesuai dengan karakteristik yang ada dinilai SLA kebersihan.

Penelitian ini menggunakan metode K-Means clustering untuk mengelompokkan stasiun berdasarkan kesamaan karakteristik yang ada dinilai SLA kebersihan. Dari penelitian yang dilakukan oleh shefia metode K-Means dianggap lebih efisien dan juga efektif sesuai yang diharapkan, lalu pada penelitian oleh Ediyanto Dalam Penelitiannya yang berjudul “Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis” dalam penelitian ini menjelaskan metode K-Means cukup efektif ditetapkan dalam proses pengklasifikasian karakteristik terhadap objek penelitian, selain K-Means juga tidak terpengaruh terhadap urutan objek yang digunakan, hal ini dibuktikan ketika peneliti mencoba menentukan secara acak titik awal pusat cluster dari salah satu objek pada permulaan perhitungan [5] didalam metode K-Means diperlukanya jumlah cluster terlebih dahulu sebelum Melakukan pengklasteran Oleh karena itu, penelitian ini juga menggunakan analisis elbow method dilakukan untuk menentukan jumlah cluster yang paling tepat dalam mengelompokkan stasiun. Dengan demikian, hasil pengelompokan diharapkan dapat lebih akurat dan relevan.

## **B. Metode**

### **1. Metode Elbow**

Metode Elbow merupakan teknik yang digunakan untuk menentukan jumlah cluster optimal dengan cara menghitung nilai Sum of Square Error (SSE) dari masing-masing cluster. Semakin besar selisih nilai SSE dengan cluster selanjutnya sehingga membentuk titik sudut siku, maka semakin baik jumlah cluster [6]. Metode ini menguji nilai SSE dengan jumlah

cluster yang berbeda-beda kemudian mencari nilai SSE dengan selisih terbesar atau yang membentuk sudut paling siku pada grafik elbow dalam menentukan jumlah cluster terbaik Pada metode Elbow nilai cluster terbaik yang akan diambil dari nilai Sum of Square Error (SSE) yang mengalami penurunan yang signifikan dan berbentuk siku saat digambarkan dalam bentuk grafik dengan rumus:

$$SSE = \sum_{K=1}^K \sum_{X_i} |X_i - C_k|^2$$

Keterangan :

K : Nilai cluster ke-i

$X_i$  : Jarak data ke-i

$C_k$  : Pusat cluster ke-k

## 2. Metode K-Means

K-Means Cluster Analysis merupakan salah satu metode cluster analysis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster atau cluster objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama diclusteran dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda diclusteran kedalam cluster yang lain [7]. Cluster-cluster yang terbentuk memiliki homogenitas internal yang tinggi dan heterogenitas eksternal yang tinggi.

Pada algoritma K-Means Cluster Analysis terdapat beberapa langkah yang harus dilakukan yaitu sebagai berikut [5]:

Tentukan nilai k sebagai jumlah cluster yang ingin dibentuk.

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

$$\underline{V}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

Keterangan:

$\underline{V}_{ij}$  : 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 variabel ke-j

Menghitung jarak antara titik centroid dengan titik objek. Untuk menghitung jarak tersebut menggunakan Euclidean Distance dengan rumus:

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

Keterangan:

$D_e$  : Euclidean Distance

$i$  : Banyaknya objek

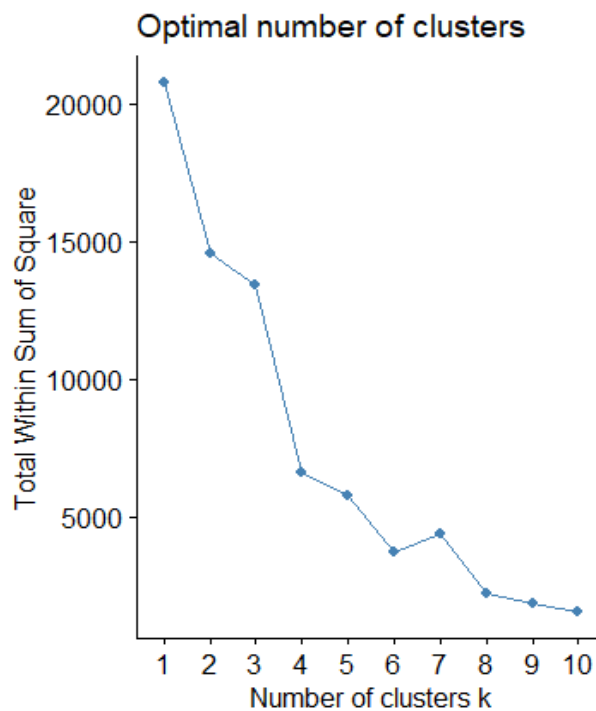
$(x,y)$  : Koordinat objek

$(s,t)$  : Koordinat centroid

Mengelompokkan objek dalam menentukan cluster dengan menghitung jarak minimum objek. Kembali ke tahap 3, melakukan perulangan hingga didapatkan nilai centroid yang dihasilkan tetap dan anggota masing-masing cluster tidak berubah. Jika kondisi ini terpenuhi maka didapatkan hasil cluster final.

### C. Hasil dan Pembahasan

Untuk menentukan jumlah cluster yang optimal dalam analisis K-Means, peneliti menggunakan metode elbow. Grafik elbow yang dihasilkan oleh RStudio menunjukkan titik belok yang mengindikasikan jumlah cluster yang paling baik.



Gambar 1. Grafik Jumlah Optimal Klaster – Metode Elbow

Berdasarkan dari gambar 1 pada grafik jumlah optimal klister menggunakan metode Elbow mengidentifikasi bahwa sudut siku-siku berada pada  $k = 4$  maka banyaknya cluster adalah 4.

Tabel 1. Data Hasil Klaster

Penamaan Cluster	Stasiun
Cluster 1	Kaliwungu, Mangkang, Gundih, Randublantung, Kapuan
Cluster 2	Jerakah, Sedadi, Jambo, Kradenan, Cepu, Larangan, Petarukan, Plabuan, Tanggung
Cluster 3	Tegal, Pemalang, Comal, Batang, Ujungnegoro, Kuripan, Weleri, Telawa, Gubug, Krengseng, Surodadi, Padas, Karangsono, Tegolawu
Cluster 4	Sulur, Sragi, Pekalongan, Klaibodri, Alastua, Brumbung, Kedungjati, Karangjati, Ngombo, Gmbringen, Dopleng, Semarang Tawang, Semarang Poncol, Wadu, Panunggalan

Tabel 1 adalah anggota dari setiap cluster, anggota paling banyak ada pada cluster 3 dengan 15 stasiun sebagai anggota. Setelah setiap cluster terbentuk maka akan ada perbedaan

dari setiap cluster atau karakteristik yang berbeda, untuk mengetahuinya dapat dilihat berdasarkan tabel berikut

**Tabel 2.** Nilai Karakteristik Klaster

iteration	Cluster			
	1	2	3	4
bahan	150,00	145,11	149,40	148,64
toilet	120,00	139,78	147,40	142,00
ruangan	150,00	148,89	149,60	140,07
halaman	150,00	120,00	121,60	136,64
peron	120,00	148,44	121,93	141,00

Berdasarkan tabel 2 dapat dilihat sebagai berikut:

1. Pada pada cluster 1 memiliki nilai paling tinggi dibandingkan pada cluster lainnya pada variabel bahan, ruangan, dan halaman dengan nilai 150,00 dan untuk variabel toilet dan peron memiliki nilai paling rendah dibandingkan cluster lainnya yaitu dengan nilai 120,00.
2. Pada cluster 2 nilai pada variabel bahan adalah 145,11 pada variabel toilet adalah 139,78 pada variabel ruangan adalah 148,89 pada variabel halaman memiliki nilai paling rendah dibandingkan cluster lainnya yaitu 120,00 dan pada variabel peron dengan nilai 148,44.
3. Pada cluster 3 nilai pada variabel bahan adalah 147,40 pada variabel toilet adalah 147,40 pada variabel ruangan adalah 149,40 pada variabel halaman adalah 121,60 dan pada variabel adalah 121,93.
4. Pada cluster 4 nilai pada variabel bahan adalah 148,64 pada variabel toilet adalah 142,00 pada variabel ruangan adalah 140,07 pada variabel halaman adalah 136,64 dan pada variabel peron adalah 141,00.

#### D. Simpulan

Berdasarkan analisis cluster menggunakan metode K-Means terhadap nilai Service Level Agreement (SLA) stasiun di wilayah Daop 4 Semarang, penelitian ini berhasil mengidentifikasi adalah: (1.) Cluster yang terbentuk ada 4 cluster dengan cluster 1 sebanyak 5 stasiun, cluster 2 sebanyak 9 stasiun, cluster 3 sebanyak 15 stasiun dan cluster 4 sebanyak 14 stasiun dengan 43 total stasiun; (2.) adapun karakteristik dari setiap stasiun pada cluster 1 lebih tinggi pada tiga variabel pada cluster 2 dan 3 dengan setiap clusternya ada pada rata-rata, dan pada cluster 4 yang tidak memiliki nilai yang rendah pada variabelnya.

#### E. Acknowledgment

Penulis menyampaikan ucapan terima kasih kepada PT KAI DAOP 4 atas kerja sama dalam menyelesaikan penelitian ini.

#### Daftar Pustaka

- [1] A. Mahiruna *et al.*, "Optimizing Medical Image Security Using Combined DWT-DCT-SVD Watermarking and RLE Compression Strategies," vol. 5, no. 1, 2024, doi: 10.26714/jichi.v5i1.14256.Optimizing.
- [2] A. A. Nuha, "Problematika Sampah dan Upaya Menjaga Kebersihan Lingkungan di

- Dusun Krajan Desa Randuagung Kecamatan Randuagung Kabupaten Lumajang,” *Khidmatuna J. Pengabd. Masy.*, vol. 1, no. 2, p. 1, 2021, doi: 10.54471/khidmatuna.v1i2.1011.
- [3] A. D and A. M. Ferro, “Evaluasi Kinerja Fasilitas Stasiun Bojonegoro Menggunakan Metode IPA (Importance Performance Analysis) dan Standar Pelayanan Minimum,” *J. Tek. Sipil*, vol. 3, no. 1, pp. 19–24, 2022, doi: 10.31284/j.jts.2022.v3i1.2411.
- [4] W. Riyanta, “Persepsi penumpang kereta api terhadap tingkat pelayanan atasiun Tugu Yogyakarta,” *J. Manaj. Dirgant.*, vol. 8, no. 9, p. 47, 2015.
- [5] R. Supardi and I. Kanedi, “Implementasi Metode Algoritma K-Means Clustering pada Toko Eidelweis,” *J. Teknol. Inf.*, vol. 4, no. 2, pp. 270–277, 2020, doi: 10.36294/jurti.v4i2.1444.
- [6] M. Sholeh and K. Aeni, “Perbandingan Evaluasi Metode Davies Bouldin, Elbow dan Silhouette pada Model Clustering dengan Menggunakan Algoritma K-Means,” *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 8, no. 1, p. 56, 2023, doi: 10.30998/string.v8i1.16388.
- [7] Ediyanto, M. N. Mara, and N. Satyahadewi, “Pengklasifikasian karakteristik dengan metode k-means cluster analysis,” *Bul. Ilm. Mat. Stat. dan Ter.*, vol. 02, no. 2, pp. 133–136, 2013.