

## Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Wilayah pada Pelanggaran Lalu Lintas

Agustiena Merdekawati<sup>1</sup>, Jefina Tri Kumalasari<sup>2</sup>

<sup>1</sup>Universitas Bina Sarana Informatika, [agustiena.atd@bsi.ac.id](mailto:agustiena.atd@bsi.ac.id), Bekasi, Indonesia

<sup>2</sup>Universitas Bina Sarana Informatika, [jefina.jtk@bsi.ac.id](mailto:jefina.jtk@bsi.ac.id), Bekasi, Indonesia

### Informasi Makalah

Submit : Apr 20, 2024  
Revisi : May 20, 2024  
Diterima : May 30, 2024

### Kata Kunci :

Clustering;  
Pelanggaran;  
K-Means;  
K-Medoids;  
*Dalvies Bouldin*.

### Abstrak

Dengan banyaknya jumlah kendaraan yang terus meningkat dapat bertambahnya angka tingkat kemacetan. Semakin meningkatnya kemacetan suatu wilayah maka dapat menyebabkan peningkatan pelanggaran, seperti pelanggaran rambu lalu lintas, tidak memakai helm, surat berkendara yang tidak lengkap, melawan arah, dan lainnya. Berbagai cara untuk menertibkan lalu lintas terus dilakukan kepolisian dan dinas perhubungan agar terciptanya kenyamanan dan keselamatan lalu lintas. Tujuan dari penelitian ini yaitu untuk menentukan pola pengelompokan data pelanggaran lalu lintas, memudahkan untuk analisa wilayah mana yang memiliki pelanggaran paling banyak sehingga dapat membuat kebijakan atau aturan untuk mengurangi pelanggaran dan tingkat kecelakaan pun menjadi menurun. Dengan menggunakan tahap *Knowledge Discovery in Database (KDD)* dalam konsep data mining sebagai metode penelitian. Penelitian ini membandingkan metode algoritma k-means dengan algoritma k-medoids, dengan pengukuran *performance* menggunakan *Dalvies Bouldin Index (DBI)*. Hasil dari penelitian ini, memiliki kesamaan kelompok *cluster* yaitu dua, dengan *cluster 0* wilayah pelanggaran lalu lintas Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah, serta *cluster 1* dengan wilayah Kemayoran. Hasil *performance* dengan menggunakan *Dalvies Bouldin Index (DBI)* pada metode k-means sebesar -0,197 dan dengan metode k-medoids sebesar -0,228. Dari hasil *performance* kedua metode tersebut, pada model k-medoids memiliki hasil lebih besar dibanding model k-means, sehingga model k-means memiliki hasil yang mendekati 0, yaitu sebesar -0,197 sehingga dapat disimpulkan K-Means memiliki hasil yang lebih baik.

### Abstract

As the number of vehicles continues to increase, the level of congestion can increase. The increasing congestion in an area can cause an increase in violations, such as violating traffic signs, not wearing a helmet, incomplete driving licenses, going against the direction, and others. The police and transportation services continue to use various methods to regulate traffic to ensure

traffic comfort and safety. The aim of this research is to determine the grouping pattern of traffic violation data, making it easier to analyze which areas have the most violations so that policies or regulations can be created to reduce violations and reduce accident rates. By using the Knowledge Discovery in Database (KDD) stage in the concept of data mining as a research method. This research compares the k-means algorithm method with the k-medoids algorithm, with performance measurement using the Dalvies Bouldin Index (DBI). The results of this research have the same cluster groups, namely two, with cluster 0 the traffic violation areas of Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, and Pal Merah, and cluster 1 with the Kemayoran area. The performance results using the Dalvis Bouldin Index (DBI) on the k-means method were -0.197 and with the k-medoids method were -0.228. From the performance results of the two methods, the k-medoids model had greater results than the k-means model, so the k-means model has results that are close to 0, namely -0.197 so it can be concluded that K-Means has better results.

## 1. Pendahuluan

Transportasi merupakan sarana yang memberikan kemudahan dalam melakukan kegiatan manusia, sebagai alat untuk memindahkan barang atau manusia dari suatu tempat ketempat lainnya menggunakan suatu kendaraan yang dapat digerakan oleh manusia atau mesin. Dengan banyaknya jumlah kendaraan yang terus meningkat dapat bertambahnya angka tingkat kemacetan. Semakin meningkatnya kemacetan suatu wilayah maka dapat menyebabkan peningkatan pelanggaran, seperti pelanggaran rambu lalu lintas, tidak memakai helm, surat berkendara yang tidak lengkap, melawan arah, dan lainnya.

Dari [www.liputan6.com](http://www.liputan6.com), terdapat 1.617.566 pelanggaran diwilayah DKI Jakarta dan sekitarnya, pada tahun 2018. Angka tersebut meningkat sekitar 24,13% dari tahun sebelumnya yang mencatat angka 1.303.157 pelanggaran (Ikbal, 2019). Menurut Brigjen Ahmad Ramadhan (Kurniawan, 2023), terdapat jumlah penindakan pelanggaran lalu lintas pada operasi patuh yang dilakukan pada tanggal 11 Juli 2023 sampai dengan 18 Juli 2023 berjumlah 29.211 dan pelanggaran dengan berupa peneguran terhadap 242.836 pengendara.

Menurut Undang-Undang No.22 Tahun 2009 tertuang berdisiplin, mulus, selamat dan

teratur dalam berlalu lintas sesuai dengan hak dan kewajiban pemakai jalan serta terhindar dari hambatan dan kemacetan di jalan. Dengan adanya tata tertib lalu lintas dapat menciptakan, berkontribusi, dan menjaga keamanan, keselamatan, keteraturan dan kelancaran lalu lintas. Salah satu bagian DKI Jakarta yang memiliki tingkat pelanggaran tinggi yaitu Jakarta Pusat.

Berbagai cara untuk menertibkan lalu lintas terus dilakukan kepolisian dan dinas perhubungan agar terciptanya kenyamanan dan keselamatan lalu lintas, dengan cara melakukan razia penilangan terhadap kelengkapan surat kendaraan, kelayakan kendaraan, serta melakukan himbauan melalui poster dan teguran.

Penggunaan piranti teknologi informasi yang tepat akan dihasilkan informasi yang tepat dan akurat sesuai dengan kebutuhan sehingga keputusan dapat diambil dengan cepat (Abdilah et al., 2018). Dengan ketersediaan informasi yang mudah didapatkan dan perkembangan teknologi maju akan semakin berkembangnya metode analisa akan lebih efisien. Pendekatan yang menggunakan algoritma dan memanfaatkan komputer, seperti:

1. Pengelompokan kondisi jalan di kota Semarang dengan menggunakan k-medoids, dengan mengelompokan data menjadi 4 *cluster* dengan kondisi jalan

- yang baik, sedang, rusak ringan dan rusak berat. (Asmiatun et al., 2020)
2. Algoritma k-means digunakan untuk mengelompokkan data pelanggaran lalu lintas sehingga dapat menentukan pelanggaran apa yang paling banyak terjadi dan memudahkan pengelompokan. Data ini dikelompokkan menjadi 2 *cluster*, yaitu banyak dan sedikit. (Wasih et al., 2022)
  3. Dengan ditemukannya tidak pidana yang semakin banyak di Indonesia, diperlukan suatu pengelompokan daerah rawan kejahatan agar mengetahui apakah suatu daerah membutuhkan pengawasan lebih atau tidak. Tindak pidana ini dikelompokkan menggunakan algoritma k-means dan k-medoids dengan mengelompokkan menjadi tingkat kriminalitas tinggi dan kriminalitas rendah. (Tampubolon et al., 2021)

RapidMiner merupakan salah satu aplikasi yang terkenal dan kompeten dalam menganalisa data, menggambarkan suatu kemungkinan atau prediksi (Putri & Putra, 2023). Dengan menggunakan RapidMiner dapat membuat pola pengelompokan data pelanggaran lalu lintas.

Penelitian ini mengelompokkan data pelanggaran lalu lintas, dengan menggunakan teknik clustering dimana record dikumpulkan dalam database berdasarkan kondisi tertentu. (Farissa et al., 2021). Dua algoritma, yaitu algoritma k-means dan algoritma k-medoids. Algoritma K-Means salah satu algoritma unsupervised learning yang digunakan untuk mengelompokkan data sehingga ditemukan suatu pola dimana memiliki centroid sebagai pusat cluster (Han et al., 2011). K-Medoids menggunakan data menjadi titik pusat yang disebut dengan medoids. Dengan menggunakan aplikasi RapidMiner untuk melihat kelompok penyebaran daerah yang sering mengalami pelanggaran kendaraan.

## 2. Metode Penelitian

Penelitian ini menerapkan metode penelitian secara kuantitatif. Metode kuantitatif merupakan metode yang menggunakan statistika atau cara lain untuk pengukuran (Nugroho & Haritanto, 2022). Metode kuantitatif didalamnya terdapat perhitungan angka.

Suatu histori data atau informasi data dan pemakaian data dalam suatu database dalam data mining disebut dengan proses *Knowledge Discovery in Database* (KDD) atau dapat disebut dengan tahap proses KDD (Arhami & Nasir, 2020). Sehingga data mining merupakan salah satu bagian dari tahap KDD.

Berikut tahap proses KDD pada penelitian ini adalah: (Arhami & Nasir, 2020)

1. Identifikasi Tujuan  
Tujuan dari penelitian ini yaitu untuk menentukan pola pengelompokan data pelanggaran lalu lintas, memudahkan untuk analisa wilayah mana yang memiliki pelanggaran paling banyak sehingga dapat membuat kebijakan atau aturan untuk mengurangi pelanggaran dan tingkat kecelakaan pun menjadi menurun.
2. Spesifikasi Permasalahan  
Permasalahan yang terjadi yaitu meningkatnya pelanggaran lalu lintas, sehingga meningkatnya kecelakaan lalu lintas di wilayah di Jakarta Pusat.
3. Seleksi Data  
Proses seleksi data disini yaitu proses memilih data yang sesuai dengan kebutuhan. Pada tahap ini memilih *field* yang dibutuhkan untuk penelitian, seperti: lokasi dan jenis pelanggaran pengendara bermotor.
4. *Preprocessing Data*  
Setelah melakukan pemilihan atribut data yang dibutuhkan, selanjutnya masuk ketahap *preprocessing*. Pada tahap *preprocessing* melakukan pembersihan data dengan menghapus data yang duplikat, menghilangkan *missing data* dan data menghapus data yang tidak konsisten

hingga normalisasi data. Tahap ini merupakan langkah yang sangat penting, karena ini sebagai langkah awal sehingga mempengaruhi hasil akhir dari pembentukan suatu pola pengelompokan. Pada tahap ini, jumlah data akan berkurang karena proses penghapusan data yang tidak atau kurang relevan.

5. Data mining

Selanjutnya melakukan proses pengolahan data mining menggunakan aplikasi *Rapid Miner* dengan menggunakan algoritma k-means dan k-medoids dengan menggunakan 9 cluster wilayah, yaitu menteng, sawah besar, tanah abang, cempaka putih, gambir, johar baru, pal merah, dan kemayoran. Hasil kedua algoritma tersebut dibandingkan dengan mengukur akurasi.

6. Pola atau Model

Hasil pengolahan data mining, yaitu menghasilkan pola pengelompokan data pelanggaran lalu lintas.

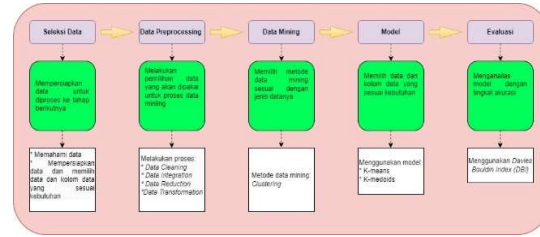
7. Interpretasi atau evaluasi

Evaluasi clustering dilakukan dengan tujuan untuk mengetahui seberapa baik kualitas dari hasil clustering

Selanjutnya melakukan pengujian pada aplikasi *Rapid Miner* dengan menggunakan *Dalvies-Bouldin Index (DBI)* untuk menghitung akurasi metode algoritma k-means dan algoritma k-medoids. Hasil dari perbandingan kedua algoritma tersebut digunakan untuk implementasi pengelompokan pelanggaran lalu lintas sehingga dapat mengurangi tingkat pelanggaran dan kecelakaan. Sehingga pihak berwenang dapat membuat aturan dan kebijakan sesuai lokasi yang memiliki tingkat pelanggaran terbanyak

3. Hasil dan Pembahasan

Kerangka penelitian pada penelitian ini berdasarkan KDD, di tunjukkan gambar 1.



Gambar 1. Kerangka Penelitian

3.1 Seleksi Data

Mengumpulkan memilih data pelanggaran. Data pelanggaran yang dipilih dari bulan januari sampai agustus tahun 2020 wilayah Jakarta Pusat dalam bentuk excel berjumlah 3073 pelanggaran. Gambar 2 contoh data pelanggaran pada penelitian ini:

JENIS PELANGGARAN	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	TEMPAT KELAJARAN PELANGGARA
	tidak ada	ada	ada	ada	ada	ada	ada	ada	ada	ada	ada	ada	ada	ada	ada	
WYBANG TRAYEK-STUK HABA HABA-BELAKU	1															DAKUH ABANG
TLANGAN STUK-STUK HABA HABA-BELAKU-TAIRA SM	1	1	1													DAKUH ATAS
WYBANG TRAYEK-TAIRA SM-APM-BELAKU		1	1	1					1							FAMORIN
TRANGAN KPS HABA HABA-BELAKU-KP HABA HABA-BELAKU TRAYEK	1	1	1	1					1							FAMORIN
STUK HABA HABA-BELAKU	1															INDOTEL
KPS HABA HABA-BELAKU-BAN GUNDUL							1									REYASURAB
TAIRA SURAT SURAT	1	1	1	1	1				1							REYASURAB
KP HABA HABA-BELAKU	1															REYASURAB
STUK HABA HABA-BELAKU	1															KS TUBUN
KP HABA HABA-BELAKU	1															KS TUBUN
STUK HABA HABA-BELAKU	1															SENIL
BODY KEROPOS-BAN GUNDUL		1						1								SHAWI SOLA
STUK HABA HABA-BELAKU	1															REYASURAB
REYASURAB									1							REYASURAB
KPS HABA HABA-BELAKU					1											SIRAGA RAG

Gambar 2. Data Pelanggaran

3.2 Data Preprocessing

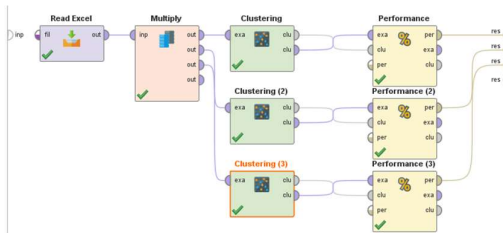
Selanjutnya dilakukan penghapusan data yang hilang (*missing value*), menghapus data yang duplikat, serta data yang tidak lengkap. Lalu data tersebut dilakukan pengelompokan sesuai dengan wilayah dan jenis pelanggaran. Dalam penelitian ini dikelompokkan menjadi 7 variabel, yaitu: pelanggaran mengenai kelengkapan surat. Pelanggaran mengenai kelengkapan kendaraan (*body keropos*, perubahan rangka, *spedometer*, lampu dan lainnya, serta seragam bagi supir kendaraan umum), pelanggaran marka rambu lalu lintas, pelanggaran melawan arah, pelanggaran untuk kendaraan yang tidak layak jalan, pelanggaran untuk kendaraan yang melebihi muatan, pelanggaran mengenai trayek (tarif dan bongkar muat barang). Gambar 3 adalah data hasil *preprocessing*.

No	Variabel	Kelengkapan Survei (X1)	Kelengkapan Kendaraan (X2)	Marka Rambu (X3)	Melampaui Arak (X4)	Tidak Layak Jalan (X5)	Mutu Q
1	Meateng	98	26	162	4	0	2
2	Sawah besar	136	50	71	0	0	3
3	Semen	76	39	147	56	0	0
4	Tanah alang	156	41	195	4	0	0
5	Cempaka putih	173	64	26	0	0	5
6	Gusuk	103	45	98	1	0	1
7	Johar baru	11	9	40	0	0	0
8	Pal merah	28	19	35	1	0	0
9	Kemayoran	453	183	174	49	4	4
	Total	1230	476	948	115	15	15

Gambar 3. Data Hasil Preprocessing

### 3.3 Data Mining dan Model

Pada tahap ini memilih jenis metode data mining *clustering* dengan metode k-means dan k-medoids. Pengolahan data mining menggunakan tools RapidMiner. Awal pengolahan dengan RapidMiner dengan melakukan pengujian *Davies Bouldin Index* (DBI), mengecek nilai DBI dengan masing-masing kelompok kluster lalu pilih yang nilainya paling kecil. Nilai DBI suatu kluster yang minimum akan ditafsirkan skema kluster tersebut bernilai optimal. Pengujian gambar 4.



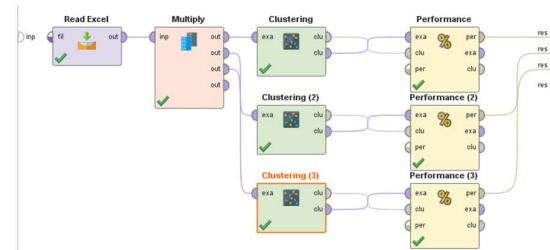
Gambar 4. Pengujian DBI pada algoritma k-means

Pada gambar 4 digunakan operator Multiply sehingga salinan objek pada RapidMiner didapatkan. Operator Performance digunakan untuk mengukur hasil kinerja dari model. Hal yang sama akan diberlakukan pada pengujian K-Medoids. Pada gambar 4 Model clustering menggunakan algoritma K-Means dengan nilai  $k = 2, k = 3, k = 4$ . Dengan hasil DBI pada tabel 1.

Cluster (K)	Hasil DBI
2	-0,197
3	-0,720
4	-0,517

Sumber : Hasil Penelitian (2023)

Berdasarkan tabel 1, nilai hasil DBI terkecil yaitu -0,197 sehingga nilai yang optimal pada kluster 2. Gambar 5, pengujian DBI menggunakan algoritma K-Medoids.



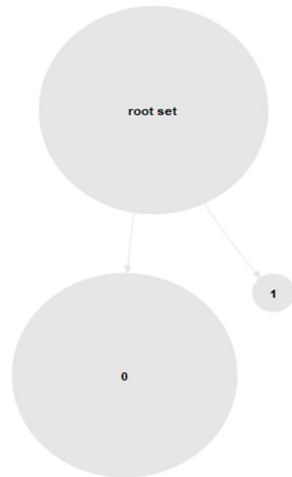
Gambar 5. Pengujian DBI pada algoritma k-medoids

Pada gambar 5, Model clustering menggunakan algoritma k-medoids dengan nilai  $k = 2, k = 3, k = 4$ . Dengan hasil DBI pada tabel 2.

Cluster (K)	Hasil DBI
2	-0,228
3	-0,589
4	-0,509

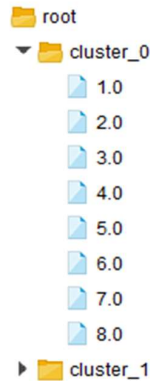
Sumber : Hasil Penelitian (2023)

Berdasarkan tabel 2, nilai hasil DBI terkecil yaitu -0,228 sehingga nilai yang optimal pada kluster 2. Selanjutnya melakukan pengelompokan atau klusterisasi dengan memilih  $k = 2$ . Karena hasil DBI pada K-Means dan K-Medoids yang terkecil pada *cluster 2*. Gambar 6 adalah hasil pengelompokan menggunakan rapidminer dengan algoritma K-Means.



Gambar 6. Clustering K-Means

Berdasarkan gambar 6, menerangkan menghasilkan 2 kelompok wilayah pelanggaran lalu lintas. Kelompok pertama dengan *cluster 0* sebanyak 8 wilayah dan kelompok kedua dengan *cluster 1* dengan 1 wilayah.



Gambar 7. Hasil Pengelompokan Wilayah Dengan Algoritma K-Means

Berdasarkan gambar 7 diatas, merupakan hasil pengelompokan wilayah di Jakarta berdasarkan kelompok klaster. Tabel 3 menunjukkan hasil pengelompokan wilayah dengan algoritma K-Means.

Tabel 3. Detail Pengelompokan Wilayah Dengan Algoritma K-Means

Clustering	Wilayah
<i>Cluster 0</i>	Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah
<i>Cluster 1</i>	Kemayoran

Sumber : Hasil Penelitian (2023)

Selanjutnya terdapat hasil titik *centroid* pada masing-masing *cluster*. Tabel 4 adalah hasil *centroid* dengan algoritma K-Means.

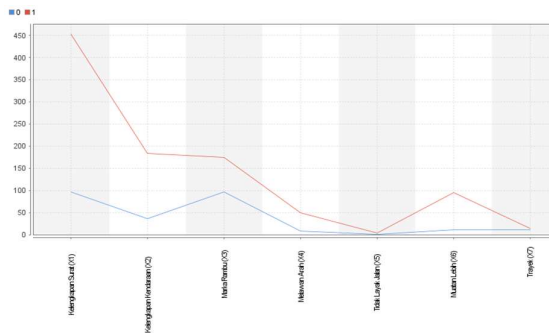
Tabel 4. Tabel Centroid K-Means

Attribute	Cluster 0	Cluster 1
Kelengkapan Surat (X1)	97,125	453
Kelengkapan Kendaraan (X2)	36,625	183
Marka Rambu (X3)	96,750	174
Melawan Arah (X4)	8,250	49
Tidak Layak Jalan (X5)	1,375	4
Muatan Lebih (X6)	10,875	95
Trayek (X7)	11,625	14

Sumber : Hasil Penelitian (2023)

Berdasarkan Tabel 4, pada pelanggaran lalu lintas mengenai kelengkapan surat (X1) nilai tertinggi pada *cluster 1* (Kemayoran) dan terendah pada *cluster 0* (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai kelengkapan kendaraan (X2) nilai tertinggi pada *cluster 1* (Kemayoran) dan terendah pada *cluster 0* (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai marka rambu (X3) nilai tertinggi pada *cluster 1* (Kemayoran) dan terendah pada *cluster 0* (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai melawan arah (X4) nilai tertinggi pada *cluster 1* (Kemayoran) dan terendah pada *cluster 0* (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru,

dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai tidak layak jalan (X5) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai muatan lebih (X6) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai trayek (X7) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Lebih jelasnya digambarkan melalui grafik pada gambar 8.



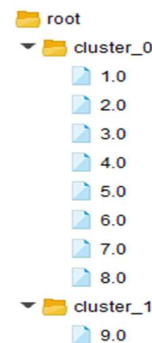
Gambar 8. Grafik *Clustering* Dengan K-Means

Gambar 9 merupakan hasil pengelompokan menggunakan rapidminer dengan algoritma K-Medoids ( $k = 2$ ).



Gambar 9. *Clustering* K-Medoids

Berdasarkan gambar 9, menerangkan menghasilkan 2 kelompok wilayah pelanggaran lalu lintas. Kelompok pertama dengan *cluster* 0 sebanyak 8 wilayah dan kelompok kedua dengan *cluster* 1 dengan 1 wilayah,



Gambar 10. Hasil Pengelompokan Wilayah Dengan Algoritma K-Medoids

Berdasarkan gambar 10, merupakan hasil pengelompokan wilayah berdasarkan kelompok klaster. Berikut hasil pengelompokan wilayah dengan algoritma K-Medoids.

Tabel 5. Detail Pengelompokan Wilayah Dengan Algoritma K-Medoids

<i>Clustering</i>	<i>Wilayah</i>
<i>Cluster</i> 0	Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih,

Clustering	Wilayah
	Gambir, Johar Baru, dan Pal Merah
Cluster 1	Kemayoran

Sumber : Hasil Penelitian (2023)

Tabel 6 menunjukkan hasil *centroid* dengan algoritma K-Medoids.

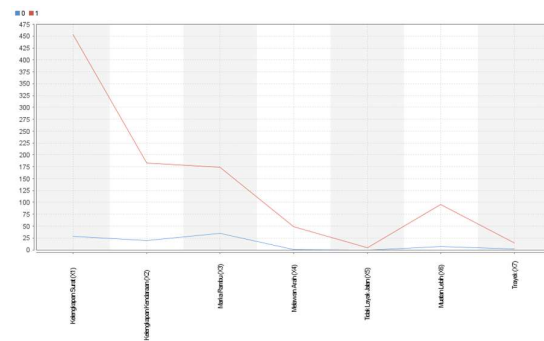
Tabel 6. Tabel Centroid K-Medoids

Attribute	Cluster 0	Cluster 1
Kelengkapan Surat (X1)	28	453
Kelengkapan Kendaraan (X2)	19	183
Marka Rambu (X3)	35	174
Melawan Arah (X4)	1	49
Tidak Layak Jalan (X5)	0	4
Muatan Lebih (X6)	7	95
Trayek (X7)	2	14

Sumber : Hasil Penelitian (2023)

Berdasarkan Tabel 6, pada pelanggaran lalu lintas mengenai kelengkapan surat (X1) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai kelengkapan kendaraan (X2) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai marka rambu (X3) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai melawan arah (X4) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai tidak layak jalan (X5) nilai

tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai muatan lebih (X6) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Tingkat pelanggaran lalu lintas mengenai trayek (X7) nilai tertinggi pada *cluster* 1 (Kemayoran) dan terendah pada *cluster* 0 (Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah). Gambar melalui grafik ditunjukkan pada gambar 11.



Gambar 11. Grafik Clustering Dengan K-Medoids

### 3.4 Evaluasi

Setelah pemodelan selesai, tahap berikutnya masuk ke tahap evaluasi. Pada evaluasi ini, peneliti melakukan analisa terhadap model yang telah menghasilkan pola. Evaluasi ini dilakukan untuk memperkuat bahwa pola yang dibuat pada masing-masing metode telah memenuhi tingkat keakuratan yang sesuai dan memberikan hasil yang bermanfaat.

Pengujian pada penelitian ini menggunakan *Davies Bouldin Index* (DBI) pada masing-masing metode dengan menggunakan  $k = 2$ . Hasil *Davies Bouldin Index* (DBI) pada metode k-means di tunjukkan pada gambar 12.



## PerformanceVector

```
PerformanceVector:  
Avg. within centroid distance: -6656.042  
Avg. within centroid distance_cluster_0: -7488.047  
Avg. within centroid distance_cluster_1: -0.000  
Davies Bouldin: -0.197
```

Gambar 12. Hasil Pengujian DBI pada K-Means

Hasil *Davies Bouldin Index* (DBI) pada metode kmedoids dtunjukkan pada gambar 12.

## PerformanceVector

```
PerformanceVector:  
Avg. within centroid distance: -14713.000  
Avg. within centroid distance_cluster_0: -16552.125  
Avg. within centroid distance_cluster_1: -0.000  
Davies Bouldin: -0.228
```

Gambar 12. Hasil Pengujian DBI pada K-Medoids

Perbandingan hasil *Davies Bouldin Index* (DBI) pada metode k-means dengan k-medoids ditunjukkan pada tabel 7.

Tabel 7. Perbandingan Hasil DBI

Metode	Davies Bouldin Index (DBI)
K-Means	-0,197
K-Medoids	-0,228

Sumber : Hasil Penelitian (2023)

Berdasarkan pada Tabel 7, hasil DBI k-means lebih kecil dibandingkan dengan hasil DBI k-medoids. Dari hasil tersebut yang memiliki nilai mendekati 0 yaitu DBI pada k-means, sehingga tingkat performancenya lebih baik dibandingkan k-medoids

## 4. Simpulan

Berdasarkan penelitian yang dilakukan dari tahap awal sampai dengan tahap evaluasi, menggunakan data dan parameter yang sesuai kebutuhan, dengan menggunakan dua metode yaitu K-Means dan K-Medoids, lalu membandingkan hasil dari model dan performance masing-masing metode, didapat

kesimpulan bahwa model kmeans memiliki performance yang lebih baik dibandingkan dengan K-Medoids, yaitu sebesar -0,197. Hasil model dengan menggunakan k-means dan K-Medoids, yaitu memberikan hasil yang sama dengan menghasilkan dua kelompok *cluster* pelanggaran lalu lintas yaitu *cluster* 0 dan *cluster* 1, serta wilayah *cluster* 0 dan *cluster* 1 pada algoritma K-Means dan K-Medoids juga sama, yaitu pada *cluster* 0 dengan pelanggaran lalu lintas pada wilayah Menteng, Sawah Besar, Senen, Tanah Abang, Cempaka Putih, Gambir, Johar Baru, dan Pal Merah, dan *cluster* 1 dengan pelanggaran lalu lintas pada wilayah Kemayoran. Dilihat dari hasil centroid pada setiap jenis pelanggaran lalu lintas, *cluster* 1 pada masing-masing metode memiliki nilai paling besar dibanding *cluster* 0.

Berdasarkan hasil ini, model K-Means dapat direkomendasikan untuk pengelompokan wilayah pelanggaran lalu lintas. Akan tetapi, pemilihan model tergantung pada tujuan dan data yang akan kita teliti yang relevan.

## 5. Referensi

- Abdilah, A., Mardiyani, E., & Nawawi, I. (2018). Aplikasi Komputer dan Smartphone Berbasis Android untuk Menangani Reservasi Hotel pada Citi Smart Hotel - BSD. *Jurnal Tehnik Komputer*, 4(2), 2. <https://doi.org/10.31294/jtk.v4i2.3597>
- Arhami, M., & Nasir, M. (2020). *DATA MINING-Algoritma dan Implementasi* (1). Andi Publishing.
- Asmiatun, S., Wakhidah, N., & Putri, A. N. (2020). Penerapan Metode K-Medoids Untuk Pengelompokan Kondisi Jalan Di Kota Semarang. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 6(2), 171–180. <https://doi.org/10.35957/jatisi.v6i2.193>
- Farissa, R. A., Mayasari, R., & Umaidah, Y. (2021). Perbandingan Algoritma K-Means dan K-Medoids Untuk Pengelompokan

- Data Obat dengan Silhouette Coefficient.  
In *Journal of Applied Informatics and Computing (JAIC)* (Vol. 5, Issue 2).  
<http://jurnal.polibatam.ac.id/index.php/JAIC>
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*.
- Ikbal, M. (2019). *Parah, Pelanggaran Lalu Lintas di DKI Jakarta Meningkatkan 24 Persen*. Liputan6.Com.  
<https://www.liputan6.com/otomotif/read/3920444/parah-pelanggaran-lalu-lintas-di-dki-jakarta-meningkat-24-persen>
- Kurniawan, D. (2023). *Operasi Patuh 2023: 29 Ribu Pelanggar Lalu Lintas Kena Tindak*. Tempo.Co.  
<https://otomotif.tempo.co/read/1750564/operasi-patuh-2023-29-ribu-pelanggar-lalu-lintas-kena-tindak>
- Nugroho, A. S., & Haritanto, W. (2022). *Metode Penelitian Kuantitatif Dengan Pendekatan Statistika (Teori, Implementasi dan Praktik dengan SPSS)* (I). CV Andi Offset.
- Putri, D. I., & Putra, M. Y. (2023). *Komparasi Algoritma Dalam Memprediksi Perubahan Harga Saham Goto Menggunakan Rapidminer*. *Jurnal Khatulistiwa Informatika*, 11(1), 14–20.
- Tampubolon, H. D., Suhada, S., Safii, M., Solikhun, & Suhendro, D. (2021). *Penerapan Algoritma K-Means dan K-Medoids Clustering untuk Mengelompokkan Tindak Kriminalitas Berdasarkan Provinsi*. *Jurnal Ilmu Komputer Dan Teknologi*, 2(2), 6–12.  
<https://doi.org/10.35960/ikomti.v2i2.703>
- Wasih, E. S., Rustam, Rahmatullah, S., Mintoro, S., & Ngajiyanto. (2022). *Implementasi Data Mining Pada Data Pelanggaran Lalu Lintas di Lampung Utara Menggunakan Algoritma K-Means (Studi Kasus Kejaksaaan Negeri Lampung Utara)*. *Escaf*, 1259–1275.