



Clustering Analysis of Call Reports for Data-Driven Response Planning

Analisis Pengelompokan Laporan Panggilan untuk Perencanaan Respons Berbasis Data

Retno Tri Cahyani^{1*}, Imam Yuadi², Hendro Margono³

¹Jurusan Pengembangan Sumber Daya Manusia, Fakultas Sekolah Pascasarjana,
Universitas Airlangga, Indonesia

^{2,3}Fakultas Ilmu Sosial dan Politik, Universitas Airlangga, Indonesia

E-Mail: ¹retno.tri.cahyani-2024@pasca.unair.ac.id,

²imam.yuadi@fisip.unair.ac.id, ³hendro.margono@fisip.unair.ac.id

Received Jul 08th 2025; Revised Aug 24th 2025; Accepted Sep 03rd 2025; Available Online Oct 30th 2025

Corresponding Author: Retno Tri Cahyani

Copyright © 2025 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

Every year, the 112 Call Center of Sidoarjo Regency receives thousands of public reports, covering various types of incidents such as fires, traffic accidents, medical emergencies, dangling cables, fallen trees, and street lighting problems. This study analyzes 6,207 reports and focuses on the geographic coordinates of incidents with the objective of clustering the spatial patterns of reports to support more responsive public service management. To achieve this goal, two clustering algorithms, K-Means and K-Medoids, were applied. The Elbow method was used to determine the number of clusters ($k=3$), where the optimum point is indicated when the inertia value begins to decrease linearly. The analysis was conducted using Google Colab and the Python programming language, supported by libraries such as scikit-learn, pyclustering, and matplotlib for modeling and visualization. The results show that K-Medoids produced more geographically structured clusters, while K-Means generated overlapping clusters. With a Silhouette Score of 0.479, higher than K-Means at only 0.193, K-Medoids proved superior in forming compact and consistent clusters. K-Medoids successfully grouped incident-prone areas (Waru, Gedangan) and infrastructure-dominant areas (Sidoarjo, Candi) into spatially appropriate clusters. This analysis also identified the characteristics of each cluster based on report types, ranging from medical emergencies to street lighting issues. These findings are valuable for supporting resource allocation and improving the efficiency of public services in the development of smart cities..

Keywords: Call Report, Clustering, Elbow Method, K-Means, K-Medoids

Abstrak

Setiap tahun, Call Center 112 Kabupaten Sidoarjo menerima ribuan laporan dari masyarakat, yang mencakup berbagai jenis kejadian seperti kebakaran, kecelakaan lalu lintas, darurat medis, kabel menjuntai, pohon tumbang, dan masalah PJU. Penelitian ini menganalisis 6.207 laporan tersebut dan berfokus pada koordinat lokasi kejadian dengan tujuan untuk mengelompokkan pola spasial laporan sehingga dapat mendukung tata Kelola pelayanan publik yang lebih responsif. Untuk mencapai tujuan tersebut digunakan dua algoritma pembelajaran yaitu K-Means dan K-Medoids. Metode Elbow digunakan untuk menentukan jumlah kluster ($k=3$). Metode ini menunjukkan titik optimum ketika nilai inerti mulai menurun secara linier. Analisis dilakukan menggunakan Google Colab dan bahasa pemrograman Python, dan ada dukungan pustaka untuk visualisasi seperti scikit-learn, pyclustering, dan matplotlib. Hasil visualisasi menunjukkan bahwa K-Medoids membentuk kluster yang lebih terstruktur secara geografis, sedangkan K-Means menghasilkan kluster yang tumpang tindih. Dengan Silhouette Score 0,479, yang lebih tinggi dari K-Means hanya 0,193, K-Medoids terbukti lebih unggul dalam membentuk kluster yang kompak dan konsisten. K-Medoids berhasil mengelompokkan wilayah yang rawan insiden (Waru, Gedangan) dan wilayah infrastruktur dominan (Sidoarjo, Candi) ke dalam kluster yang sesuai secara spasial. Analisis ini mengidentifikasi fitur tiap kluster berdasarkan jenis laporan, mulai dari darurat medis hingga masalah PJU. Penemuan ini berguna untuk mendukung alokasi sumber daya dan layanan publik yang lebih efisien saat membangun kota pintar.

Kata Kunci : Clustering, Elbow Method, K-Means, K-Medoids, Laporan Panggilan

1. PENDAHULUAN

Di era digital saat ini, semakin banyak orang yang membutuhkan layanan publik yang efisien dan responsif. Untuk memenuhi kebutuhan ini, pengambilan keputusan memerlukan penggunaan teknologi berbasis data. Panggilan darurat 112 adalah salah satu layanan penting yang sangat penting bagi kehidupan masyarakat karena menawarkan jalur komunikasi langsung antara orang-orang dan fasilitas penanganan darurat untuk berbagai jenis kejadian, seperti kebakaran, kecelakaan lalu lintas, gangguan di fasilitas umum, dan kondisi medis darurat. Layanan ini telah digunakan secara aktif di beberapa daerah di Indonesia, salah satunya adalah Kabupaten Sidoarjo. Layanan *Call Center* 112 Kabupaten Sidoarjo telah menerima 6.207 laporan dari masyarakat sejak didirikan. Laporan tersebut berasal dari berbagai kategori seperti kebakaran, laka lantas, darurat medis, pohon tumbang, banjir, dan padamnya PJU (Penerangan Jalan Umum). Namun, jumlah laporan yang banyak dan tersebar luas sering menyulitkan pemerintah daerah untuk membuat rencana penanganan yang tepat dan efektif. Akibatnya, untuk memudahkan proses perencanaan dan alokasi sumber daya penanganan, diperlukan pendekatan analisis data yang dapat mengelompokkan laporan berdasarkan pola geografisnya. *Clustering*, atau pembagian data ke dalam kelompok berdasarkan kesamaan tertentu, adalah salah satu metode yang dapat digunakan. Karena kemampuan untuk menemukan pola-pola tersembunyi dalam data tanpa perlu adanya label sebelumnya, *clustering* sangat populer dalam data mining dan analitik spasial [1].

Studi ini melihat bagaimana data laporan *Call Center* 112 di Kabupaten Sidoarjo dikelompokkan berdasarkan *latitude* dan *longitude*. Dua algoritma *clustering* yang digunakan adalah *K-Means* dan *K-Medoids*. *K-Means* telah menjadi salah satu metode *clustering* yang paling populer karena mampu menangani data berskala besar [2]. Meskipun demikian, ada kelemahan metode ini terhadap *outlier* [2]. Sebuah penelitian baru oleh Rudianto [3] menemukan bahwa meskipun *K-Means* dapat memproses data dengan cepat, kepekaannya terhadap nilai ekstrim dapat menyebabkan klasterisasi data yang tidak akurat. Karena mereka lebih tahan terhadap *outlier* dan suara, *K-Medoids* digunakan sebagai metode pembanding. *K-Medoids* digunakan sebagai metode pembanding karena lebih tahan terhadap *outlier* dan *noise*. Studi oleh Rohman dan Wibowo [4] menemukan bahwa, karena menggunakan titik medoid sebagai pusat kluster daripada rata-rata data, *K-Medoids* memiliki keunggulan dalam menghasilkan kluster yang lebih stabil pada data spasial.

Karena sangat efektif dalam menemukan pola data dan mengelompokkan peristiwa yang serupa, metode *clustering* telah banyak digunakan dalam analisis spasial dan layanan darurat. Karena efisiensi dan kesederhanaannya, *K-Means* adalah salah satu algoritma *clustering* yang paling populer [2]. Namun, karena sensitivitasnya terhadap *outlier* dan asumsi bentuk *cluster* yang bulat, beberapa penelitian mempertimbangkan metode alternatif seperti *K-Medoids*. Sebuah penelitian yang membandingkan *K-Means* dan *K-Medoids* dalam analisis data kecelakaan lalu lintas [3] menemukan bahwa *K-Medoids* lebih baik dalam menangani data dengan suara. Studi serupa juga ditemukan dalam penelitian Rohman dan Wibowo [4] yang menggunakan kedua algoritma untuk klasterisasi lokasi ke data kecelakaan lalu lintas.

Selanjutnya, penelitian [5] menilai kedua algoritma dalam perencanaan tanggap darurat medis dan menemukan bahwa *K-Medoids* menghasilkan ukuran kluster yang lebih seimbang, yang sangat penting untuk distribusi sumber daya yang adil. Selain itu, penelitian [6] menggunakan metode elbow untuk menemukan jumlah kluster yang ideal dan menemukan bahwa validitas kluster meningkat saat menggabungkan jenis kejadian dan koordinat lokasi. Namun, pada data lalu lintas perkotaan, *K-Means* lebih baik untuk data skala besar, tetapi *K-Medoids* lebih baik dalam situasi distribusi data yang tidak beraturan. Selain itu, studi [7] memeriksa metode *clustering* hibrida yang menggabungkan *K-Means* dan pendekatan berbasis densitas. Namun, mereka tetap mengakui bahwa algoritma partisi tradisional seperti *K-Means* dan *K-Medoids* masih merupakan dasar penting untuk analisis data spasio-temporal. Penelitian ini memberikan bukti kuat bahwa analisis komparatif algoritma ini sangat penting untuk pengelompokan data laporan panggilan darurat publik seperti yang dilakukan dalam penelitian.

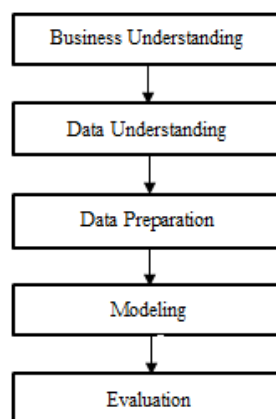
Kebutuhan akan sistem perencanaan respons yang lebih efisien dan berbasis data untuk laporan kedaruratan masyarakat mendorong penelitian ini. Dengan demikian, pertanyaan penelitian yang diajukan adalah sebagai berikut: Bagaimana kedua algoritma *K-Means* dan *K-Medoids* digunakan untuk mengelompokkan laporan *Call Center* 112 Kabupaten Sidoarjo berdasarkan koordinat lokasi (*latitude* dan *longitude*)? Dan algoritma mana yang memberikan hasil *clustering* berdasarkan skor silhouette yang paling optimal? Penelitian ini bertujuan untuk mengelompokkan laporan masyarakat melalui *Call Center* 112 Kabupaten Sidoarjo berdasarkan koordinat geografis menggunakan *K-Means* dan *K-Medoids* dengan *Elbow method* untuk menentukan jumlah optimal *cluster*. Kemudian untuk membandingkan performa kedua algoritma digunakan *Silhouette Score*. Penelitian ini berkontribusi untuk memberikan pemetaan pola spasial laporan darurat yang dapat mendukung perencanaan respons berbasis data, sekaligus untuk menyediakan analisis komparatif yang dapat menjadi acuan pemilihan algoritma *clustering* dalam pengelolaan laporan kedaruratan di tingkat daerah.

2. BAHAN DAN METODE

Ada empat metode penelitian yang umum digunakan, yaitu penelitian tindakan, eksperimen, studi kasus, dan survei. Dalam penelitian, metode yang digunakan mengacu pada bagaimana masalah yang diangkat dalam penelitian dapat diselesaikan. Seperti halnya pengumpulan data, merumuskan hipotesis, menguji hipotesis, menginterpretasikan hasil hingga membuat kesimpulan [8]. Pada penelitian ini menggunakan penelitian eksperimen dimana peneliti menguji kebenaran suatu hipotesis dengan statistik dan menghubungkannya dengan masalah penelitian [9]. Dua jenis penelitian eksperimen adalah eksperimen absolut dan komparatif. Eksperimen absolut menghasilkan dampak, seperti pengaruh Kepuasan Kerja dan deskripsi pekerjaan terhadap retensi karyawan. Sedangkan eksperimen komparatif, di sisi lain, melibatkan perbandingan dua objek, seperti membandingkan dua algoritma untuk menentukan hasil statistik mana yang lebih baik [9]. Dan dalam penelitian ini, jenis penelitian yang digunakan adalah eksperimen komparatif. Rancangan penelitian ini adalah membandingkan dan mengevaluasi dua metode seperti *K-Means Cluster* dan *K-Medoids*. Untuk mengetahui model mana yang memiliki akurasi yang lebih tinggi dalam mengelompokkan data laporan *Call Center* 112 berdasarkan lokasi geografis.

Metode pengumpulan data dibagi menjadi dua yaitu primer dan sekunder. Data primer merupakan data yang dikumpulkan oleh peneliti secara langsung dari sumber pertama. Pengumpulan data primer biasa dilakukan dengan metode yang dirancang secara khusus untuk menjawab pertanyaan penelitian atau menguji hipotesis. Data primer meliputi wawancara, observasi dan *survey* atau kuesioner [10]. Sedangkan metode pengumpulan data sekunder yaitu pengambilan data dari data yang sudah dipublikasikan sebelumnya oleh pihak lain seperti Lembaga pemerintah, organisasi, atau peneliti sebelumnya. Data sekunder meliputi Dokumen resmi, Literatur akademik dan Sumber Media atau *Online* [11]. Pada penelitian ini menggunakan data sekunder yaitu data laporan masyarakat yang ada di Kabupaten Sidoarjo serta bersumber dari buku, jurnal atau literatur terdahulu.

Cross-Industry Standard Process for Data Mining (CRISP-DM) digunakan dalam penelitian ini. Menurut [12] CRISP-DM dalam penelitian ini dibagi dalam lima tahap.



Gambar 1. Tahapan Penelitian

2.1 Business Understanding

Data laporan *Call Center* 112 Kabupaten Sidoarjo dikelompokkan berdasarkan lokasi geografis (*latitude* dan *longitude*). Mengidentifikasi pola penyebaran lokasi kejadian darurat untuk mendukung perencanaan respons berbasis data secara efektif adalah tujuan utama dari proses *clustering* ini. Selain itu, penelitian ini membandingkan dua algoritma *clustering*, *K-Means* dan *K-Medoids*, berdasarkan akurasi hasil *clustering* yang dihasilkan menggunakan nilai *Silhouette Score*. Hasil yang diharapkan adalah mengetahui potensi algoritma pembelajaran mesin dalam melakukan *clustering* berdasarkan letak *latitude* dan *longitude* suatu wilayah, yang berkontribusi pada peningkatan respons dan alokasi sumber daya efisien di *smart city*.

2.2 Data Understanding

Data ini diambil dari 6.207 laporan masyarakat ke *Call Center* 112 Kabupaten Sidoarjo. Kategori laporan, waktu, dan koordinat lokasi (*latitude* dan *longitude*) dimasukkan ke dalam tiap laporan. Karena fokus penelitian adalah pada sebaran geografis laporan kejadian, hanya atribut *latitude* dan *longitude* yang digunakan untuk proses *clustering*.

Tabel 1. Data Laporan *Call Center* 112 Kabupaten Sidoarjo

Waktu_Lapor	Kategori	Latitude	Longitude	Kecamatan
2022-06-21 15:19:23	LAKA LANTAS	-7.3860486	112.6094576	KRIAN

Waktu_Lapor	Kategori	Latitude	Longitude	Kecamatan
2022-06-22 00:43:12	KAMTIBMAS	-7.389686	112.7279293	GEDANGAN
2022-06-21 20:55:23	TRAFFIC LIGHT BERMASALAH	-7.4471138	112.7015698	SIDOARJO
2022-06-22 00:42:12	KAMTIBMAS	-7.4470261	112.6557377	WONOAYU
2022-06-22 12:10:25	TRAFFIC LIGHT BERMASALAH	-7.441567	112.683095	SIDOARJO
2022-06-22 16:36:55	LAKA LANTAS	-7.3512451	112.6917401	TAMAN
2022-06-23 15:19:28	LAKA LANTAS	-7.3670223	112.729057	WARU
2022-06-24 07:10:00	TRAFFIC LIGHT BERMASALAH	-7.4107679	112.578586	KRIAN
2022-06-24 10:51:16	BANJIR LUAPAN SUNGAI	-7.3972848	112.608878	KRIAN

2.3 Data Preparation

Pada tahap ini, langkah-langkah pra-pemrosesan data berikut dilakukan:

1. Seleksi kolom: untuk analisis, hanya kolom *latitude* dan *longitude* yang dipilih.
2. Pembersihan data adalah proses mengevaluasi data *outlier* dan nilai kosong (*null*) yang dapat mempengaruhi hasil *clustering*.
3. Penentuan jumlah *cluster* : penentuan jumlah *cluster* dilakukan dengan *Elbow method* untuk menentukan jumlah *cluster* (*k*) yang ideal berdasarkan nilai inerti. *Elbow Method* merupakan metode yang melihat persentase variansi yang dijelaskan sebagai fungsi dari jumlah *cluster* [13]. Dari hasil *Elbow method* didapatkan 3 *cluster* sebagai nilai optimal.

2.4 Modeling

Pemodelan merupakan tahapan untuk mengolah data dengan model algoritma yang telah ditentukan untuk melakukan *clustering*. Dalam permodelan, penelitian ini menggunakan algoritma *K-Means* dan *K-Medoids* untuk membentuk kelompok berdasarkan koordinat geografis dengan pembagian *cluster* menjadi 3 *cluster* untuk mengukur kinerja model.

2.4.1 K-Means

K-Means merupakan algoritma *unsupervised learning* yang digunakan untuk membagi data ke dalam *k cluster* berdasarkan kesamaan lokasi titik data [14]. Dimulai dari memilih centroid awal secara acak dan kemudian menetapkan setiap titik data ke *cluster* terdekat berdasarkan jarak geometris. Pusat *cluster* tidak berubah secara signifikan setelah *centroid* diperbarui secara *iterative* hingga konvergensi [15]. Keunggulan dari *K-Means* adalah efisiensi komputasionalnya dan kemampuan untuk menangani *dataset* yang sangat besar, namun saat rentan terhadap *outlier* [16]. *K-Means* sering digunakan dalam berbagai bidang seperti segmentasi pelanggan pengolahan citra dan juga analisis spasial [16].

2.4.2 K-Medoids

K-Medoids adalah jenis *clustering* partisi yang berbeda yang memilih pusat *cluster* sebagai konep medoid, yaitu salah satu titik nyata dalam data, sehingga lebih tahan terhadap *outlier* daripada *K-Means* [17]. Algoritma umum *K-Medoids*, adalah PAM, yang meskipun kuat namun memiliki kompleksitas tinggi [17]. Studi terbaru, seperti BanditPAM++, telah mempercepat proses ini secara signifikan. Ini menawarkan metode yang lebih efektif yang kompleksitasnya mendekati $O(k)$ lebih cepat daripada algoritma awal [18]. Selain itu, relevansi *K-Medoids* dalam konteks spasial dan *non-linear* ditunjukkan oleh aplikasi kontemporer seperti *clustering* pelanggan PAM di Indonesia [19].

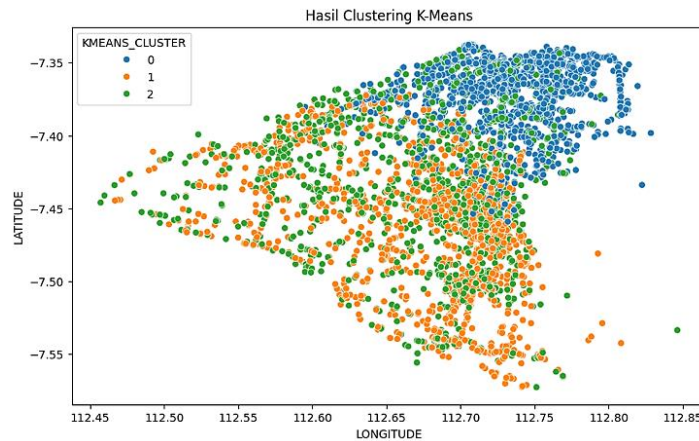
2.5 Evaluation

Untuk menilai model, nilai *Silhouette Score* dihitung untuk masing-masing metode. Hal ini digunakan untuk mempelajari dan memahami jarak pemisah antara *cluster* yang dihasilkan [20]. Hasil dengan nilai *Silhouette Score* lebih tinggi menunjukkan pemisahan *cluster* yang lebih baik dan kohesi internal *cluster* yang lebih baik.

3. HASIL DAN PEMBAHASAN

Penelitian ini dilakukan dengan menggunakan Google Colab dengan Bahasa python. Algoritma yang digunakan untuk melakukan *clustering* yaitu *K-Means Cluster* dan *K-Medoids*. Dengan membagi *cluster* menjadi 3 bagian dan mendapatkan hasil sebagai berikut :

Dengan menggunakan algoritma *K-Means* berdasarkan hasil visualisasi pada Gambar 2, *clustering* menghasilkan tiga kelompok utama yang didasarkan pada lokasi geografis dalam laporan masyarakat. Klaster tidak membentuk batas wilayah yang kaku, dan terjadi overlap atau pencampuran antar klaster, terutama di bagian tengah peta. Ini menunjukkan karakter *K-Means*, yang dibangun berdasarkan perhitungan rata-rata posisi, yang rentan terhadap pengaruh *outlier* dan distribusi data yang tidak simetris.



Gambar 2. Hasil *Clustering K-Means*

Tabel 2. Sebaran *Cluster K-Means* Berdasarkan Kecamatan

Kecamatan	Cluster		
	0	1	2
BALONG BENDO	0	43	46
BUDURAN	140	142	195
CANDI	3	262	167
GEDANGAN	335	21	129
JABON	0	101	35
KREMBUNG	0	62	38
KRIAN	15	121	171
PORONG	0	118	43
PRAMBON	0	75	76
SEDATI	268	1	42
SIDOARJO	132	477	350
SUKODONO	109	92	179
TAMAN	368	36	164
TANGGULANGIN	0	176	144
TARIK	0	57	55
TULANGAN	0	82	56
WARU	764	0	83
WONOAYU	1	101	131

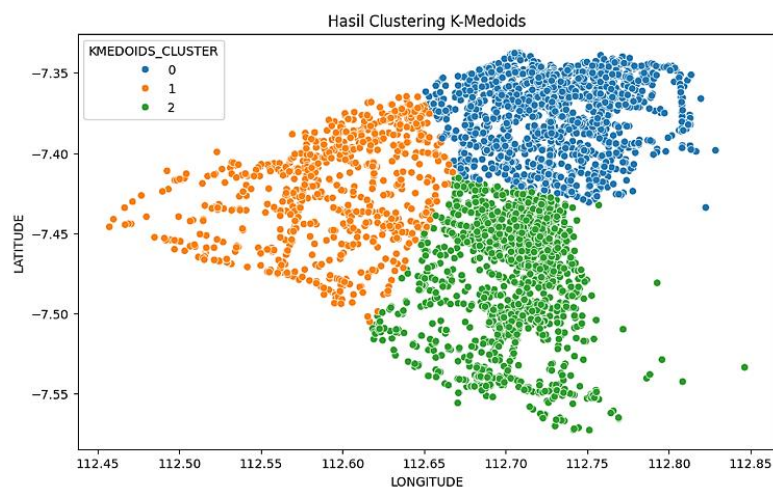
Pada analisis Tabel 2 berdasarkan kecamatan, menunjukkan bahwa *Cluster 0* di dominasi oleh kecamatan Waru dengan 764 laporan, Taman 368, Gedangan 335 laporan dan Sedati 268 laporan. Wilayah ini berada di sisi barat laut hingga timur laut Kabupaten Sidoarjo. *Cluster 1* di dominasi oleh kecamatan Sidoarjo dengan 477 laporan, Candi 262 laporan dan Tanggulangin dengan 176 laporan. Pada *cluster* ini menunjukkan keragaman jenis laporan. Dan di *Cluster 2*, terdapat 350 laporan di Sidoarjo, 195 di Buduran dan 179 di Kecamatan Sukodono. *Cluster* ini menunjukkan laporan yang sedang hingga tinggi.

Tabel 3. Sebaran *Cluster K-Means* Berdasarkan Kategori

Kategori	Cluster		
	0	1	2
ADMINISTRASI KECAMATAN	12	12	16
BANJIR LUAPAN SUNGAI	25	15	25
BEASISWA	0	0	1
BENCANA ALAM	4	2	9
DARURAT MEDIS	227	102	116
EVAKUASI HEWAN	130	117	102
EVAKUASI LAIN	9	2	3
EVAKUASI MANUSIA	8	2	2
JALAN RUSAK	131	86	119
KABEL MENJUNTAI	17	25	21
KAMTIBMAS	70	75	53
KEBAKARAN	148	106	125
KECELAKAAN AIR	0	1	2
KEMACETAN	7	14	15

Kategori	Cluster		
	0	1	2
KRIMINALITAS	144	97	30
LAKA LANTAS	443	143	203
ODGJ+PMKS	55	40	43
PARKIR LIAR	2	5	4
PENANGANAN SAMPAH	19	13	11
PENELANTARAN/KEKERASAN	13	9	6
PEREMPUAN/ANAK			
PENEMUAN JENAZAH	3	6	2
PENEMUAN ORANG HILANG	1	1	2
PERBAIKAN RAMBU LANTAS	1	0	3
PERMASALAHAN PDAM	125	61	73
PERMASALAHAN PJU	372	852	909
PERMASALAHAN REKLAME	3	6	2
PERMASALAHAN SUNGAI	18	17	23
PERMASALAHAN TIANG	4	2	3
PERMINTAAN DEREK	34	20	22
POHON TUMBANG	36	43	63
POLUSI INDUSTRI	1	3	5
TRAFFIC LIGHT BERMASALAH	69	88	84
TUMPAHAN OLI	2	0	3
WABAH ULAT BULU	3	2	4

Dan pada *clustering* berdasarkan kategori laporan dalam Tabel 3, dapat dilihat bahwa *cluster* 0 di dominasi oleh laporan laka lintas sebanyak 443 laporan, Darurat medis 227 laporan dan kebakaran sebanyak 148 laporan. Hal ini menunjukkan bahwa daerah tersebut rawan dengan kecelakaan darurat medis dan juga kebakaran. Untuk *Cluster* 1 lebih menunjukkan keberagaman yaitu mencakup laporan evakuasi hewan, laka lintas, kebakaran dan gangguan ketertiban umum. Sedangkan untuk *cluster* 2 sangat menonjol dalam laporan permasalahan penerangan jalan umum dengan 909 laporan. Permasalahan Perusahaan Daerah Air Minum (PDAM) dan Permasalahan Sungai.



Gambar 3. Hasil *Clustering K-Medoids*

Gambar 3 menunjukkan hasil *Clustering K-Medoids*. untuk hasil visualisasi spasial dari *Clustering K-Medoids*. memperlihatkan batas geografis klaster yang lebih terstruktur dan bersih. Area tersebut dibagi menjadi tiga bagian besar yang membentang dari barat ke timur secara hampir simetris. Ini karena *K-Medoids* lebih tahan terhadap *outlier* karena menggunakan titik medoid aktual sebagai pusat klaster daripada centroid rata-rata dalam data.

Jika dilihat pada Tabel 4. Dalam *Cluster* 0 mencakup daerah Waru (846 laporan), Gedangan (485), Taman (436), Sedati (310), Sukodono (283), dan sebagian Buduran (220). Kecamatan-kecamatan ini, yang sebagian besar terletak di wilayah barat laut hingga tengah Kabupaten Sidoarjo, dikenal memiliki kepadatan penduduk yang tinggi dan banyak aktivitas masyarakat yang berkembang.

Sedangkan pada *cluster* 1 ini mencakup beberapa daerah seperti Krian (307 laporan), Wonoayu (206), Balong Bendo (89), Tarik (112), Prambon (151), dan sebagian kecil Sukodono dan Tulangan. Wilayah ini cenderung merupakan zona transisi atau semi-perdesaan, dengan kepadatan sedang dan pelaporan insiden tidak seaktif klaster 0.

Dan untuk *cluster* 2, Klaster ini terdiri dari Sidoarjo (950 laporan), Candi (431), Tanggulangin (320), Buduran (257), Porong (161), Jabon (136), dan sebagian kecil Krembung dan Tulangan. Ini adalah klaster dengan jumlah laporan infrastruktur tertinggi. Laporan paling banyak berkaitan dengan kabel menjuntai, pohon tumbang, PDAM, traffic light bermasalah, dan PJU (940 laporan)

Tabel 4. Sebaran *Cluster K-Medoids* Berdasarkan Kecamatan

Kecamatan	<i>Cluster</i>		
	0	1	2
BALONG BENDO	0	89	0
BUDURAN	220	0	257
CANDI	0	1	431
GEDANGAN	485	0	0
JABON	0	0	136
KREMBUNG	0	9	91
KRIAN	0	307	0
PORONG	0	0	161
PRAMBON	0	151	0
SEDATI	310	0	1
SIDOARJO	7	2	950
SUKODONO	283	53	44
TAMAN	436	132	0
TANGGULANGIN	0	0	320
TARIK	0	112	0
TULANGAN	0	49	89
WARU	846	0	2
WONOAYU	0	206	27

Tabel 5. Sebaran *Cluster K-Medoids* Berdasarkan Kategori

Kategori	<i>Cluster</i>		
	0	1	2
ADMINISTRASI	16	9	15
KECAMATAN			
BANJIR LUAPAN SUNGAI	27	15	23
BEASISWA	1	0	0
BENCANA ALAM	4	4	7
DARURAT MEDIS	251	39	155
EVAKUASI HEWAN	159	48	142
EVAKUASI LAIN	6	3	5
EVAKUASI MANUSIA	7	1	4
JALAN RUSAK	163	80	93
KABEL MENJUNTAI	22	8	33
KAMTIBMAS	50	18	130
KEBAKARAN	186	58	135
KECELAKAAN AIR	0	1	2
KEMACETAN	19	8	9
KRIMINALITAS	88	41	142
LAKA LANTAS	487	103	199
ODGJ+PMKS	72	9	57
PARKIR LIAR	2	0	9
PENANGANAN SAMPAH	23	2	18
PENELANTARAN/KEKERAS			
AN PEREMPUAN/ANAK	15	2	11
PENEMUAN JENAZAH	3	2	6
PENEMUAN ORANG			
HILANG	3	0	1
PERBAIKAN RAMBU			
LANTAS	2	0	2
PERMASALAHAN PDAM	144	10	105
PERMASALAHAN PJU	629	564	940
PERMASALAHAN			
REKLAME	4	2	5
PERMASALAHAN SUNGAI	25	9	24
PERMASALAHAN TIANG	5	1	3
PERMINTAAN DEREK	40	11	25
POHON TUMBANG	45	23	74

Kategori	Cluster		
	0	1	2
POLUSI INDUSTRI	2	4	3
TRAFFIC LIGHT	82	32	127
BERMASALAH			
TUMPAHAN OLI	2	2	1
WABAH ULAT BULU	3	2	4

Untuk sebaran *Cluster K-Medoids* berdasarkan Kategori terlihat pada Tabel 5. Terdapat sebaran laporan yang terbagi dalam 3 *cluster*. Jenis laporan yang paling banyak ditemukan dalam *cluster* 0 adalah laka lintas (487 laporan), darurat medis (251 laporan), keamanan dan ketertiban umum, dan permasalahan PJU (629 laporan), yang menunjukkan bahwa wilayah dalam *cluster* ini memiliki tingkat risiko tinggi terhadap insiden darurat, yang membutuhkan penanganan cepat dan ketersediaan layanan publik yang intensif.

Kemudian pada *Cluster* 1, *Cluster* ini menerima laporan yang lebih merata daripada dua klaster lainnya, tetapi volumenya lebih sedikit. Tidak banyak laporan tentang laka lintas, darurat medis, dan evakuasi hewan. Klaster ini dapat digambarkan sebagai area yang menghubungkan pusat aktivitas dengan wilayah pinggiran. Ini memerlukan pendekatan pemerataan layanan daripada intensifikasi. Terakhir pada *Cluster* 2, dimana pada *Cluster* ini dapat dianggap sebagai area yang berfokus pada pemeliharaan dan penguatan fasilitas umum karena banyaknya laporan terkait infrastruktur teknis. Ini dapat menunjukkan betapa padatnya infrastruktur yang membutuhkan perawatan rutin dan betapa banyaknya penanganan yang membutuhkan perhatian pemerintah.

Jika dilihat dari kategori laporan: *Cluster* 0 adalah wilayah dengan dominasi insiden darurat dan keamanan, seperti laka lintas, darurat medis, dan kriminalitas, sekaligus menghadapi tantangan infrastruktur seperti lampu jalan dan jalan rusak. *Cluster* 1 memiliki jumlah laporan yang relatif kecil di semua kategori, menandakan wilayah dengan aktivitas sosial yang sedang dan risiko insiden rendah. *Cluster* 2 adalah *Cluster* dengan dominasi laporan infrastruktur publik, seperti PJU, PDAM, *traffic light*, dan kabel menjuntai, mencerminkan wilayah dengan konsentrasi infrastruktur tinggi atau kualitas layanan teknis yang masih kurang optimal.

Selanjutnya untuk evaluasi model menggunakan Silhouette Score. Salah satu metrik evaluasi yang digunakan untuk menilai kualitas hasil *clustering* adalah *Silhouette Score*. Skor ini menunjukkan seberapa mirip data dengan *Clusternya* dibandingkan dengan *Cluster* lain. Ini adalah interpretasi dari skor yang berkisar dari -1 hingga 1. Pada hasil *Silhouette Score* menunjukkan bahwa *K-Medoids* memiliki hasil lebih baik dengan akurasi 0.479 dibandingkan dengan *K-Means* yang hanya memiliki akurasi 0.193.

4. SIMPULAN

Pada penelitian ini mengevaluasi kinerja algoritma *K-Means Cluster* dan juga *K-Medoids*. Pada hasil menunjukkan bahwa algoritma *K-Medoids* lebih baik dalam melakukan segmentasi spasial laporan masyarakat dibandingkan dengan *K-Means* dalam melakukan segmentasi spasial laporan masyarakat.. Hasil visualisasi spasial mereka menunjukkan pembagian wilayah yang lebih terstruktur dan sedikit tumpang tindih antar klaster, yang didukung oleh nilai *Silhouette Score* sebesar 0.479 yang jauh lebih tinggi dibandingkan dengan *K-Means Cluster* yang hanya mempunyai *Silhouette Score* sebesar 0.193. Hal ini menunjukkan bahwa *K-Medoids* lebih tahan terhadap *outlier* dan distribusi data yang tidak simetris. yang menunjukkan bahwa algoritma *K-Medoids* menghasilkan *cluster* yang lebih stabil, terpisah, dan reversible.

Dalam pembagian *cluster*, hasil *clustering* menunjukkan tiga pola utama laporan masyarakat. *Cluster* 0 utamanya ada di kecamatan Waru, Gedangan dan Taman. Dengan dominasi insiden darurat seperti laka lintas, kriminalitas dan juga darurat medis. Hal tersebut mencerminkan daerah yang memiliki kebutuhan akan layanan yang cepat. Sementara itu *Cluster* 1 merupakan wilayah transisi dengan intensitas laporan sedang. Dan untuk *Cluster* 2 yang meliputi Sidoarjo, Candi, Tanggulangin dan juga porong cenderung memiliki masalah infrastruktur teknis yang tinggi seperti PJU padam, PDAM yang menjadi prioritas dalam hal pemeliharaan fasilitas umum. Hasilnya menunjukkan bahwa algoritma *K-Medoids* sangat efektif untuk membantu pengambilan keputusan berbasis wilayah dalam perencanaan pelayanan publik dan tanggap darurat. Hasil *cluster* dapat digunakan oleh pemerintah daerah untuk membuat rencana distribusi sumber daya. Strategi ini dapat mencakup pemeliharaan fasilitas teknis, penempatan armada layanan darurat, dan membuat kebijakan berbasis zona risiko. Sistem manajemen kota pintar atau smart city yang responsif dan efisien dapat didukung ke depan dengan integrasi data spasial dan kategorikal seperti ini.

REFERENSI

- [1] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to data mining*. Pearson Education India, 2016.
- [2] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern recognition letters*, vol. 31, no. 8, pp. 651-666, 2010.

-
- [3] R. D. Rudianto and A. W. Wijayanto, "Analisis Perbandingan K-Means dan K-Medoids dalam Pengelompokan Provinsi Berdasarkan Indeks Demokrasi Indonesia 2021," *Komputika: Jurnal Sistem Komputer*, vol. 13, no. 1, pp. 19-26, 2024.
 - [4] N. Rohman and A. Wibowo, "Perbandingan Metode K-Medoids dan Metode K-Means Dalam Analisis Segmentasi Pelanggan Mall," *SINTECH (Science and Information Technology) Journal*, vol. 7, no. 1, pp. 49-58, 2024.
 - [5] Q.-X. Zhu, X.-W. Wang, N. Zhang, Y. Xu, and Y.-L. He, "Novel K-medoids based SMOTE integrated with locality preserving projections for fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-8, 2022.
 - [6] Y. Kim and Y. Kim, "Global regionalization of heat environment quality perception based on K-means clustering and Google trends data," *Sustainable Cities and Society*, vol. 96, p. 104710, 2023.
 - [7] S. Wang, J. G. Yabes, and C.-C. H. Chang, "Hybrid density-and partition-based clustering algorithm for data with mixed-type variables," *Journal of Data Science*, vol. 19, no. 1, pp. 15-36, 2021.
 - [8] C. W. Dawson, *Projects in computing and information systems: a student's guide*. Pearson Education, 2005.
 - [9] C. R. Kothari, *Research methodology: Methods and techniques*. New Age International, 2004.
 - [10] J. W. Creswell, "Research designs. Qualitative, quantitative, and mixed methods approaches," 2009.
 - [11] A. Bryman, *Social research methods*. Oxford university press, 2016.
 - [12] P. Chapman *et al.*, "CRISP-DM 1.0: Step-by-step data mining guide," *SPSS inc*, vol. 9, no. 13, pp. 1-73, 2000.
 - [13] P. Bholowalia and A. Kumar, "EBK-means: A clustering technique based on elbow method and k-means in WSN," *International Journal of Computer Applications*, vol. 105, no. 9, 2014.
 - [14] P. Sharma, "The most comprehensive guide to k-means clustering you'll ever need," URL: <https://www.analyticsvidhya.com/blog/2019/08/comprehensiveguide-k-means-clustering>, 2019.
 - [15] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 881-892, 2002.
 - [16] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Information Sciences*, vol. 622, pp. 178-210, 2023.
 - [17] E. Schubert and P. J. Rousseeuw, "Fast and eager k-medoids clustering: O (k) runtime improvement of the PAM, CLARA, and CLARANS algorithms," *Information Systems*, vol. 101, p. 101804, 2021.
 - [18] M. Tiwari, R. Kang, D. Lee, S. Thrun, I. Shomorony, and M. J. Zhang, "BanditPAM++: Faster \$ k \$-medoids Clustering," *Advances in Neural Information Processing Systems*, vol. 36, pp. 73371-73382, 2023.
 - [19] N. Almazar and M. B. Yel, "Application of the K-Medoids Algorithm in Clustering PAM Customers Based on Provinces in Indonesia. 4," ed: April, 2024.
 - [20] G. Ogbuabor and F. Ugwoke, "Clustering algorithm for a healthcare dataset using silhouette score value," *Int. J. Comput. Sci. Inf. Technol.*, vol. 10, no. 2, pp. 27-37, 2018.
-