

Klasterisasi Pola Nilai Impor Migas Bulanan Berdasarkan Pelabuhan Bongkar di Indonesia Tahun 2024 Menggunakan Algoritma K-Means

Akram Farrasanto¹, Muhamad Raehan², Muhammad Verdy Hasan Alhafiz³

¹²³Universitas Pelita Bangsa, Jawa Barat, Indonesia

akram.farrasanto09@mhs.pelitabangsa.ac.id,

muhamad.312210266@mhs.pelitabangsa.ac.id,

verdyhasan.312210214@mhs.pelitabangsa.ac.id

ABSTRACT

Indonesia's oil and gas (O&G) distribution relies on imports through various unloading ports with different monthly patterns. This study aims to cluster O&G ports in 2024 based on monthly import values using the K-Means algorithm. The method follows Knowledge Discovery in Databases (KDD) stages: data selection, preprocessing, transformation, clustering, evaluation, and visualization. Analysis was conducted in Google Colab using Python with Scikit-learn, Pandas, and Matplotlib. Results show three main clusters: ports with high, medium, and low import volumes. Evaluation using Elbow Method and Silhouette Score confirmed that three clusters offer optimal separation. PCA visualization clearly shows cluster distribution. These findings support more efficient energy logistics planning and port infrastructure development based on data-driven insights.

Keywords: clustering, distribution, energy, K-Means, port

ABSTRAK

Distribusi minyak dan gas (migas) nasional bergantung pada aktivitas impor melalui pelabuhan bongkar di berbagai wilayah Indonesia. Tiap pelabuhan memiliki pola nilai impor yang bervariasi setiap bulan, sehingga perlu diklasterisasi untuk memahami karakteristik distribusi. Penelitian ini bertujuan mengelompokkan pelabuhan bongkar migas tahun 2024 berdasarkan nilai impor bulanan menggunakan algoritma K-Means. Metode yang digunakan mengikuti tahapan Knowledge Discovery in Databases (KDD), mulai dari pemilihan, pra-pemrosesan, transformasi data, klasterisasi, evaluasi, hingga visualisasi. Analisis dilakukan di Google Colab menggunakan Python dengan pustaka Scikit-learn, Pandas, dan Matplotlib. Hasil klasterisasi menghasilkan tiga kelompok utama: pelabuhan dengan volume impor tinggi, sedang, dan rendah. Evaluasi menggunakan Elbow Method dan Silhouette Score menunjukkan bahwa tiga klaster memberikan pemisahan optimal. Visualisasi dengan Principal Component Analysis (PCA) menunjukkan persebaran klaster yang jelas. Temuan ini memberikan gambaran yang berguna untuk mendukung perencanaan distribusi migas yang lebih efisien dan penguatan infrastruktur pelabuhan sesuai karakteristik impornya.

Kata kunci: distribusi, energi, K-Means, klasterisasi, pelabuhan

PENDAHULUAN

Indonesia sebagai negara kepulauan sangat bergantung pada impor migas untuk memenuhi kebutuhan energi. Data BPS mencatat nilai impor migas mencapai US\$2,65 miliar pada Agustus 2024, turun dari US\$3,56 miliar di Juli, lalu kembali naik menjadi US\$3,67 miliar pada Oktober. Fluktuasi ini dipengaruhi oleh harga global, infrastruktur pelabuhan, dan kebijakan nasional. Informasi bulanan per pelabuhan tersedia dalam publikasi resmi BPS. [1]

Karena setiap pelabuhan memiliki kapasitas dan karakteristik tersendiri, klusterisasi berbasis algoritma K-Means dapat digunakan untuk mengelompokkan pola impor migas bulanan. Pendekatan ini memungkinkan identifikasi pelabuhan dengan profil pengimporan serupa dan memfasilitasi perencanaan distribusi yang lebih efisien. K-Means dipilih karena praktis, cepat, dan banyak digunakan dalam analisis spasial serta temporal, seperti pada studi pengelompokan provinsi berdasarkan data COVID-19. [2]

Berdasarkan konteks tersebut, penelitian ini merumuskan permasalahan utama mengenai bagaimana pola nilai impor migas bulanan dapat diklasifikasikan secara optimal berdasarkan pelabuhan bongkar di Indonesia sepanjang tahun 2024. Secara khusus, penelitian ini berupaya menjawab tiga pertanyaan: pertama, apakah terdapat pola kluster khusus pada nilai impor migas bulanan berdasarkan pelabuhan bongkar; kedua, berapa jumlah kluster optimal yang dapat merepresentasikan pola distribusi tersebut; dan ketiga, bagaimana hasil klusterisasi ini dapat menjadi acuan dalam pengelolaan distribusi migas nasional yang lebih efisien dan tepat sasaran. Sejalan dengan itu, tujuan dari penelitian ini adalah untuk mengelompokkan pelabuhan bongkar migas di Indonesia berdasarkan pola nilai impor migas bulanan sepanjang tahun 2024 dengan menggunakan algoritma K-Means.[3]

METODE PENELITIAN

Pendekatan dan Platform

Penelitian ini menggunakan pendekatan kuantitatif berbasis data mining, dengan fokus pada algoritma klusterisasi K-Means. Analisis dilakukan secara interaktif menggunakan Google Colab, memanfaatkan pustaka Python seperti pandas, numpy, scikit-learn, serta matplotlib dan seaborn untuk visualisasi.

Sumber dan Jenis Data

Dataset yang digunakan berupa data sekunder dari publikasi BPS 2024 berjudul Nilai Impor Migas Bulanan Menurut Pelabuhan Bongkar Tahun 2024. Data mencakup nilai impor (ribuan USD) per bulan (Januari–Desember 2024) untuk masing-masing pelabuhan bongkar di Indonesia. Setiap baris merepresentasikan satu pelabuhan, dengan kolom berisi nilai bulanan serta total tahunan.[1]

Proses Knowledge Discovery in Databases (KDD)

Tahap awal analisis data dimulai dengan seleksi data untuk menyaring baris yang relevan. Selanjutnya, nilai numerik dikonversi ke tipe float, nilai hilang diimputasi dengan median, dan data dinormalisasi menggunakan StandardScaler

guna menyamakan skala antar fitur[4]. Data yang telah diproses kemudian ditransformasikan menjadi matriks fitur berukuran $m \times 12$, dengan m adalah jumlah pelabuhan dan 12 kolom mewakili nilai impor migas bulanan.[5]

Proses Klasterisasi K-Means

Setelah data diproses melalui tahap pra-pemrosesan dan transformasi, langkah selanjutnya adalah menerapkan algoritma K-Means clustering. Klasterisasi ini bertujuan untuk mengidentifikasi kelompok pelabuhan dengan karakteristik distribusi migas yang mirip, sehingga dapat mendukung perumusan strategi distribusi dan logistik yang lebih efisien. Langkah-langkah dalam algoritma K-Means adalah sebagai berikut:[6]

Menentukan jumlah klaster (k)

Pada langkah pertama, kita perlu menentukan jumlah klaster (k) yang optimal menggunakan metode Elbow, yang akan dihitung menggunakan sum of squared errors (SSE). Jumlah klaster ini dapat ditentukan dengan menganalisis hasil grafik Elbow, yang mengindikasikan titik optimal di mana penurunan SSE mulai melambat.[7]

Inisialisasi centroid

Setelah menentukan jumlah klaster (k), centroid (titik pusat) untuk masing-masing klaster akan diinisialisasi secara acak. Titik pusat ini nantinya akan digunakan untuk mengelompokkan data.

Perhitungan jarak (Euclidean Distance)

Proses klasterisasi dimulai dengan menghitung jarak antara setiap titik data dan masing-masing centroid. Rumus yang digunakan adalah Euclidean Distance sebagai berikut:

$$D(x_i, y_i) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

$D(x_i, y_i)$: adalah jarak antara titik data x_i dan pusat klaster (centroid)

y_i ,

x_i : adalah nilai data pada dimensi ke- i ,

y_i : adalah nilai pusat klaster pada dimensi ke- i ,

n : adalah jumlah dimensi atau atribut data.

Penugasan kluster

Setelah menghitung jarak, setiap data akan dikelompokkan ke dalam kluster dengan jarak terdekat dari centroid-nya. Data tersebut akan dimasukkan dalam kluster yang memiliki pusat terdekat.

Pembaharuan centroid

Setelah penugasan data ke kluster, posisi centroid akan diperbarui. Centroid baru dihitung berdasarkan rata-rata nilai dari semua data dalam kluster tersebut.

$$C_j = \frac{1}{|S_j|} \sum_{x_i \in S_j} x_i \quad (2)$$

- C_j : adalah posisi baru dari centroid kluster j,
 S_j : adalah set data dalam kluster j,
 $|S_j|$: adalah jumlah data dalam kluster j,
 x_i : adalah nilai data dalam kluster j .

Iterasi

Langkah-langkah ini diulang hingga posisi centroid tidak berubah lagi, yang menandakan bahwa proses klusterisasi telah selesai.[8]

Evaluasi Hasil Clustering

Evaluasi dilakukan menggunakan Silhouette Coefficient untuk menilai kualitas kluster yang terbentuk. Nilai berkisar dari -1 hingga +1, di mana nilai mendekati +1 menunjukkan kluster yang terpisah dengan baik [9]. Perhitungan dilakukan di Google Colab menggunakan pustaka scikit-learn, dan hasilnya digunakan untuk memvalidasi jumlah kluster optimal.

Visualisasi Hasil

Visualisasi hasil klusterisasi dilakukan menggunakan Google Colab dengan bantuan pustaka Python seperti matplotlib, seaborn, dan scikit-learn. Teknik Principal Component Analysis (PCA). Setiap titik mewakili pelabuhan bongkar dan diberi warna sesuai kluster hasil K-Means[10]. Visualisasi ini memudahkan interpretasi sebaran dan pemisahan antar kluster berdasarkan pola nilai impor migas bulanan.

Diagram Alur Metode Penelitian

Berikut adalah diagram alur yang menggambarkan langkah-langkah yang dilakukan dalam penelitian ini:



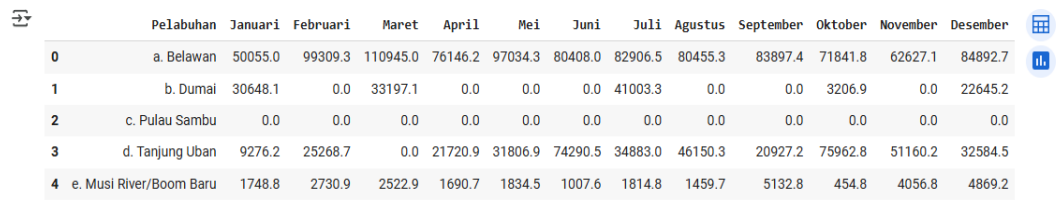
Gambar 1 Alur Metode Penelitian

HASIL DAN PEMBAHASAN

Pada bagian ini, akan dibahas mengenai hasil klusterisasi pola impor migas bulanan berdasarkan pelabuhan bongkar di Indonesia pada tahun 2024 menggunakan algoritma K-Means. Proses analisis dilakukan melalui beberapa tahapan, yaitu pemilihan data, pre-processing, transformasi data, klusterisasi, dan evaluasi hasil. Hasil yang diperoleh akan dianalisis untuk memberikan gambaran yang lebih jelas tentang distribusi impor migas di pelabuhan-pelabuhan di Indonesia.

Pemilihan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder dari Badan Pusat Statistik (BPS) tahun 2024 yang mencatat nilai impor migas bulanan berdasarkan pelabuhan bongkar di Indonesia. Data ini mencakup 41 pelabuhan setelah dilakukan pembersihan, yaitu dengan menghapus entri agregat wilayah seperti "SUMATERA:" atau "KALIMANTAN:" yang tidak mewakili pelabuhan individu. Masing-masing pelabuhan memiliki data nilai impor migas dalam satuan ribu USD untuk 12 bulan, dari Januari hingga Desember.




	Pelabuhan	Januari	Februari	Maret	April	Mei	Juni	Juli	Agustus	September	Oktober	November	Desember
0	a. Belawan	50055.0	99309.3	110945.0	76146.2	97034.3	80408.0	82906.5	80455.3	83897.4	71841.8	62627.1	84892.7
1	b. Dumai	30648.1	0.0	33197.1	0.0	0.0	0.0	41003.3	0.0	0.0	3206.9	0.0	22645.2
2	c. Pulau Sambu	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	d. Tanjung Uban	9276.2	25268.7	0.0	21720.9	31806.9	74290.5	34883.0	46150.3	20927.2	75962.8	51160.2	32584.5
4	e. Musi River/Boom Baru	1748.8	2730.9	2522.9	1690.7	1834.5	1007.6	1814.8	1459.7	5132.8	454.8	4056.8	4869.2

Gambar 2 Contoh 5 Pelabuhan Bongkar dan Nilai Impornya per Bulan

Pre-processing Data

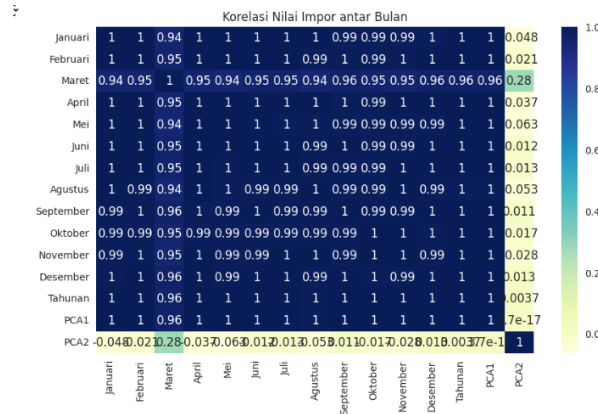
Pra-pemrosesan data dilakukan untuk menyiapkan dataset sebelum memasuki tahap klusterisasi. Tahapan ini mencakup pengubahan seluruh nilai numerik ke dalam tipe data float agar dapat diproses secara matematis, serta penanganan nilai hilang (missing values) dengan metode imputasi menggunakan nilai median. Setelah itu, dilakukan normalisasi data menggunakan StandardScaler untuk memastikan bahwa setiap fitur berada pada skala yang setara. Langkah ini penting karena algoritma K-Means sangat bergantung pada perhitungan jarak antar data, yang dapat terdistorsi apabila skala antar fitur tidak seragam.



	Januari	Februari	Maret	April	Mei	Juni	Juli	Agustus	September	Oktober	November	Desember
0	-0.325629	-0.257243	-0.264211	-0.293239	-0.244600	-0.300383	-0.305793	-0.270114	-0.259751	-0.324341	-0.299501	-0.296759
1	-0.359865	-0.417124	-0.379411	-0.415019	-0.411817	-0.418791	-0.362375	-0.415149	-0.420021	-0.414492	-0.416088	-0.388305
2	-0.413933	-0.417124	-0.428600	-0.415019	-0.411817	-0.418791	-0.417742	-0.415149	-0.420021	-0.418705	-0.416088	-0.421609
3	-0.397568	-0.376443	-0.428600	-0.380281	-0.357005	-0.309392	-0.370639	-0.331955	-0.380043	-0.318928	-0.320848	-0.373688
4	-0.410848	-0.412728	-0.424862	-0.412315	-0.408656	-0.417307	-0.415291	-0.412518	-0.410216	-0.418107	-0.408536	-0.414448

Gambar 3 Ringkasan Statistik Data yang Telah Dinormalisasi (Mean \approx 0, Std \approx 1)

Gambar di atas menunjukkan bahwa tidak ada missing values yang ditemukan dalam dataset, yang berarti data siap untuk diproses lebih lanjut.



Gambar 4 Heatmap Korelasi antar Bulan

Transformasi Data

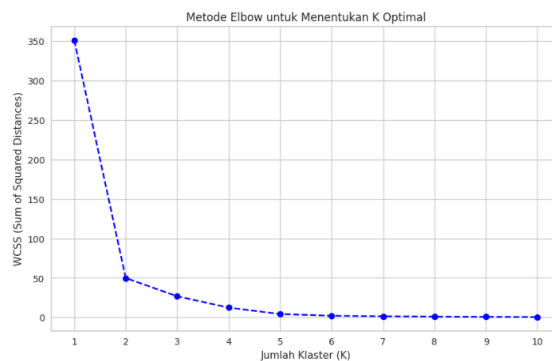
Setelah dinormalisasi, data diubah ke format array numerik berdimensi **(41 × 12)**. Setiap baris merepresentasikan pelabuhan dan kolom mewakili nilai impor migas per bulan. Format ini diperlukan untuk proses K-Means dan PCA. Data siap diproses lebih lanjut untuk penentuan jumlah kluster optimal.

Penentuan Jumlah Kluster Optimal (K)

Menentukan jumlah kluster yang optimal merupakan langkah krusial dalam proses klusterisasi menggunakan algoritma K-Means. Jumlah kluster (k) yang tidak tepat dapat menyebabkan informasi yang salah atau kurang bermanfaat. Dalam penelitian ini, dua pendekatan digunakan untuk menentukan nilai k terbaik, yaitu Metode Elbow dan Silhouette Score.

Metode Elbow untuk Menentukan Nilai k Optimal

Metode Elbow bertujuan mengidentifikasi titik optimal di mana penambahan jumlah kluster tidak lagi menghasilkan penurunan signifikan dalam nilai Within-Cluster Sum of Squares (WCSS). Dalam metode ini, nilai WCSS dihitung untuk setiap k dari 1 hingga 10. Hasilnya diplot dalam grafik, dan titik “siku” pada grafik menunjukkan nilai k optimal.



Gambar 5 Grafik Metode Elbow untuk Menentukan Nilai k Optimal

Silhouette Score

Sebagai pelengkap evaluasi, digunakan Silhouette Score untuk mengukur kualitas klasterisasi dari sisi kepadatan dan pemisahan antar klaster. Meskipun nilai tertinggi ditemukan pada $k = 2$ (0.8742), namun secara visualisasi dan konteks kebijakan, $k = 3$ dianggap lebih informatif dalam mendeskripsikan pola impor migas yang kompleks di Indonesia.

Tabel 1 Nilai Silhouette Score untuk $k = 2$ sampai $k = 10$

Jumlah Klaster	Silhouette Score
2	0.8742
3	0.6805
4	0.6481
5	0.6115
6	0.606
7	0.5708
8	0.5448
9	0.5081
10	0.4538

Implementasi K-Means dan Analisis Hasil

Implementasi K-Means dilakukan setelah data nilai impor migas bulanan dinormalisasi. Algoritma ini mengelompokkan pelabuhan bongkar di Indonesia tahun 2024 ke dalam tiga klaster berdasarkan kemiripan pola impornya. Setiap pelabuhan diberi label klaster, yang menunjukkan identitas pola impornya. Hasil awal menunjukkan mayoritas pelabuhan berada dalam klaster berimpor rendah, sementara hanya sedikit yang berimpor tinggi. Analisis lanjutan terhadap rata-rata nilai impor dan jumlah anggota tiap klaster memberikan wawasan strategis untuk pengelolaan distribusi migas nasional.

Implementasi K-Means dengan $K = 3$

Klasterisasi dilakukan menggunakan algoritma K-Means dengan jumlah klaster tiga. Dataset yang telah dinormalisasi diproses menggunakan KMeans dari scikit-learn, menghasilkan klasifikasi tiap pelabuhan bongkar tahun 2024 ke dalam salah satu dari tiga klaster berdasarkan pola impor bulannya. Label klaster ditambahkan ke dataset akhir (`df_final`). Tabel berikut menampilkan lima pelabuhan pertama beserta nilai normalisasi dan klasternya:

Tabel 2 Contoh 5 Pelabuhan Pertama dengan Label Kluster

	Pelabuhan				
	Belawan	Dumai	Pulau Sambu	Tanjung Uban	Musi River/Boom Baru
Januari	-0.39247	-0.42399	-0.47378	-0.45871	-0.47094
Februari	-0.32531	-0.47554	-0.47554	-0.43732	-0.47141
Maret	-0.32993	-0.43762	-0.4836	-0.4836	-0.4801
April	-0.35608	-0.46826	-0.46826	-0.43626	-0.46577
Mei	-0.31377	-0.4674	-0.4674	-0.41704	-0.4645
Juni	-0.36668	-0.47751	-0.47751	-0.37511	-0.47612
Juli	-0.36951	-0.42253	-0.47442	-0.43028	-0.47212

Tabel 3 Nilai Silhouette Score untuk k = 2 sampai k = 10

Jumlah Kluster	Silhouette Score
2	0.8742
3	0.6805
4	0.6481
5	0.6115
6	0.606
7	0.5708
8	0.5448
9	0.5081
10	0.4538

Implementasi K-Means dan Analisis Hasil

Implementasi K-Means dilakukan setelah data nilai impor migas bulanan dinormalisasi. Algoritma ini mengelompokkan pelabuhan bongkar di Indonesia tahun 2024 ke dalam tiga kluster berdasarkan kemiripan pola impornya. Setiap pelabuhan diberi label kluster, yang menunjukkan identitas pola impornya. Hasil awal menunjukkan mayoritas pelabuhan berada dalam kluster berimpor rendah, sementara hanya sedikit yang berimpor tinggi. Analisis lanjutan terhadap rata-rata nilai impor dan jumlah anggota tiap kluster memberikan wawasan strategis untuk pengelolaan distribusi migas nasional.

Implementasi K-Means dengan K = 3

Klasterisasi dilakukan menggunakan algoritma K-Means dengan jumlah kluster tiga. Dataset yang telah dinormalisasi diproses menggunakan KMeans dari scikit-learn, menghasilkan klasifikasi tiap pelabuhan bongkar tahun 2024 ke dalam salah satu dari tiga kluster berdasarkan pola impor bulannya. Label kluster ditambahkan ke dataset akhir (*df_final*). Tabel berikut menampilkan lima pelabuhan pertama beserta nilai normalisasi dan klasternya:

Tabel 4 Contoh 5 Pelabuhan Pertama dengan Label Kluster

	Pelabuhan				
	Belawan	Dumai	Pulau Sambu	Tanjung Uban	Musi River/Boom Baru
Januari	-0.39247	-0.42399	-0.47378	-0.45871	-0.47094
Februari	-0.32531	-0.47554	-0.47554	-0.43732	-0.47141
Maret	-0.32993	-0.43762	-0.4836	-0.4836	-0.4801
April	-0.35608	-0.46826	-0.46826	-0.43626	-0.46577
Mei	-0.31377	-0.4674	-0.4674	-0.41704	-0.4645
Juni	-0.36668	-0.47751	-0.47751	-0.37511	-0.47612
Juli	-0.36951	-0.42253	-0.47442	-0.43028	-0.47212
Agustus	-0.34183	-0.47764	-0.47764	-0.39973	-0.47517
September	-0.32747	-0.47531	-0.47531	-0.43843	-0.46626
Oktober	-0.39367	-0.47999	-0.48403	-0.38849	-0.48345
November	-0.36171	-0.4718	-0.4718	-0.38186	-0.46467
Desember	-0.36518	-0.45101	-0.48223	-0.4373	-0.47552

Analisis Karakteristik Kluster (Rata-rata Fitur per Kluster)

Setelah pelabelan kluster selesai, langkah selanjutnya adalah menganalisis karakteristik masing-masing kluster. Analisis dilakukan dengan menghitung rata-rata nilai impor migas bulanan (dalam skala asli) untuk tiap-tiap kluster. Ini membantu memahami pola distribusi impor migas dari masing-masing kelompok pelabuhan.

Tabel 5 Rata-rata Nilai Impor Bulanan per Kluster

	Klaster 0	Klaster 1	Klaster 2
Januari	631764.2	2305260	62237.38
Februari	647041.5	2472760	72772.9
Maret	400336.1	2753934	123538.1
April	641094.9	2549665	70881.43
Mei	634211.8	2363867	60915.42
Juni	704788.4	2725478	81295.1
Juli	756773.2	2960465	87784.56
Agustus	552097.3	2211998	70884.36
September	532691.3	2145597	63344.05
Oktober	774835.2	2935331	99791.93
November	515966.4	2114821	66767.82
Desember	674183.7	2733464	88820.51
Total	7465784	30272640	949033.6

Dari data tersebut, dapat disimpulkan bahwa:

- a. Kluster 1 mencerminkan pelabuhan besar dengan volume impor migas sangat tinggi secara konsisten setiap bulan.

dari pelabuhan-pelabuhan dengan volume impor tinggi secara konsisten, klaster kedua berisi pelabuhan dengan volume impor sedang, sementara klaster ketiga mencakup mayoritas pelabuhan yang memiliki volume impor rendah dan fluktuatif. Analisis karakteristik klaster ini memberikan gambaran yang lebih terstruktur mengenai peran masing-masing pelabuhan dalam sistem distribusi migas nasional. Temuan ini diharapkan dapat menjadi acuan dalam perencanaan kebijakan energi dan logistik yang lebih terarah, terutama dalam hal penguatan infrastruktur dan efisiensi distribusi migas di wilayah-wilayah strategis.

Saran

Penelitian ini dapat dikembangkan lebih lanjut dengan menambahkan variabel-variabel eksternal seperti jenis migas (misalnya BBM, LPG, dan gas alam), volume fisik dalam satuan liter atau ton, serta faktor ekonomi dan lingkungan seperti harga migas dunia dan kondisi cuaca pelabuhan. Hal ini akan memperkaya analisis dan meningkatkan akurasi klasterisasi. Selain itu, bagi pemerintah dan pengambil kebijakan, hasil klasterisasi ini dapat dijadikan sebagai dasar dalam merancang strategi distribusi energi nasional yang lebih efisien, terutama dalam mengidentifikasi pelabuhan-pelabuhan strategis untuk optimalisasi logistik migas. Dari sisi metodologi, penggunaan algoritma lain seperti DBSCAN atau hierarchical clustering juga disarankan untuk membandingkan kualitas hasil pengelompokan dan mendapatkan perspektif yang lebih luas. Terakhir, untuk mendukung visualisasi spasial yang lebih informatif, penggunaan perangkat lunak seperti Quantum GIS (QGIS) dapat menjadi pelengkap dalam menganalisis sebaran pelabuhan berdasarkan klaster secara geografis.

DAFTAR PUSTAKA

- Abdullah, D., S. Susilo., A. S. Ahmar., R. Rusli., and R. Hidayat. (2022). The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data. *Qual Quant*, 56(3), 1283-1291. doi: 10.1007/s11135-021-01176-w. [2]
- Afidah, N. Nur. (2023). Penerapan Metode Clustering dengan Algoritma K-means untuk Pengelompokan Data Migrasi Penduduk Tiap Kecamatan di Kabupaten Rembang. *PRISMA, Prosiding Seminar Nasional Matematika*, 6, 729-738, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>. [4]
- Badan Pusat Statistik. Badan Pusat Statistik. Nilai Impor Migas Bulanan Menurut Pelabuhan Bongkar, [1]
- Bagus, I., *et al.* (n.d.). Penerapan Data Mining untuk Clustering Peminat Layanan Iconnet Berdasarkan Wilayah Area Bali Menggunakan Metode K-Means. [3]
- Br Sembiring, S. Sarnauli., A. Hadi Nasyuha., I. Mariami., P. Studi Sistem Informasi, S. Triguna Dharma., and P. Studi Sistem Informasi Dosen Pembimbing. (2020). Pemanfaatan Algoritma K-means dalam Pengelompokan Data Siswa Berdasarkan Prestasi Akademik sebagai Penentuan Penerima Beasiswa di

- SMK Negeri 1 Lubuk Pakam. *Jurnal CyberTech*, 3(9).
<https://ojs.trigunadharma.ac.id/>. [6]
- Maesaroh, M., T. Nur Padilah., and J. Haerul Jaman. (2023). Penerapan Algoritma K-Means Clustering pada Pengelompokan Daerah Penyebaran Diare di Provinsi Jawa Barat. [9]
- Negara, H. A., A. R. Putra., and U. Enri. (2021). Clustering Data Ekspor Buah-Buahan Berdasarkan Negara Tujuan Menggunakan Algoritma K-Means. 7(2). [Online]. Available: <http://ejournal.fikom-unasman.ac.id>. [10]
- Riyono, J., and C. E. Pujiastuti. (2022). Simulation Of The K-Means Clustering Algorithm With The Elbow Method in Making Clusters Of Provincial Poverty Levels in Indonesia. *Jurnal Matematika MANTIK*, 8(2), 113-123. doi: 10.15642/mantik.2022.8.2.113-123. [5]
- Siregar, A. Hidayah., D. Sihotang, B., Angga Wijaya., and S. Dohot Siregar. (2024). Implementasi Algoritma K-Means Menggunakan RapidMiner untuk Klasterisasi Data Obat pada Rumah Sakit Royal Prima. [8]
- Vania, P., and B. Nurina Sari. (2023). Perbandingan Metode Elbow dan Silhouette untuk Penentuan Jumlah Klaster yang Optimal pada Clustering Produksi Padi menggunakan Algoritma K-Means. *Jurnal Ilmiah Wahana Pendidikan*, 9(21), 547-558. doi: 10.5281/zenodo.10081332. [7]