

K-Means Analysis of West Java Coastal Capture Fish Production to Increase Fisheries Productivity

Siti Salwa Azzahra ^{1*)}, Fatia Amalia Maresti ²⁾, Kiki Mustaqim ³⁾

¹⁾²⁾³⁾Prodi Sains Data, Universitas Logistik dan Bisnis Internasional

^{*)}Correspondence author: salwazahra879@gmail.com, Bandung, Indonesia

DOI: <https://doi.org/10.37012/jtik.v12i1.3508>

Abstract

Marine capture fisheries in West Java play a vital role in regional economic development, yet disparities in productivity among coastal areas remain unresolved. This study aims to classify coastal districts and cities in West Java based on key production indicators: fish catch volume, number of vessels, number of marine fishers, average weighted price per kilogram, and the number of fishing households. A K-Means Clustering approach was applied to group these regions into homogeneous clusters, enabling more targeted policy development. The analysis identified three distinct clusters: moderately productive areas with low efficiency, economically efficient areas with high-value catch but limited volume, and a dominant production zone with significantly higher outputs than others. These findings provide a clearer understanding of regional fishery dynamics and offer recommendations for each cluster. Proposed strategies include expanding infrastructure and training for low-performing regions, and focusing on product diversification, post-harvest processing, and sustainable management in high-performing areas. This study demonstrates the usefulness of data-driven methods to support marine fishery development and regional planning.

Keywords: Marine Capture Fisheries, K-Means Clustering, Regional Disparity

Abstrak

Perikanan tangkap laut di Jawa Barat memiliki peran penting dalam pembangunan ekonomi daerah, namun masih terdapat kesenjangan produktivitas antar wilayah pesisir yang belum terselesaikan. Penelitian ini bertujuan untuk mengelompokkan kabupaten dan kota pesisir di Jawa Barat berdasarkan indikator-indikator produksi utama: volume hasil tangkapan ikan, jumlah kapal, jumlah nelayan laut, harga rata-rata tertimbang per kilogram, dan jumlah rumah tangga perikanan. Pendekatan K-Means Clustering digunakan untuk mengelompokkan wilayah-wilayah ini ke dalam klaster yang memiliki karakteristik serupa, sehingga dapat mendukung perumusan kebijakan yang lebih tepat sasaran. Analisis menghasilkan tiga klaster utama: wilayah dengan produksi sedang namun efisiensi rendah, wilayah dengan efisiensi tinggi dan nilai jual ikan tinggi meskipun volume tangkapan terbatas, serta satu wilayah dominan dengan produksi yang jauh lebih tinggi dibanding wilayah lain. Temuan ini memberikan pemahaman yang lebih jelas tentang dinamika perikanan regional dan menghasilkan rekomendasi bagi masing-masing klaster. Strategi yang diusulkan meliputi peningkatan infrastruktur dan pelatihan untuk wilayah dengan kinerja rendah, serta fokus pada diversifikasi produk, pengolahan pascapanen, dan pengelolaan berkelanjutan untuk wilayah dengan kinerja tinggi. Studi ini menunjukkan pentingnya pendekatan berbasis data dalam mendukung pengembangan perikanan laut dan perencanaan wilayah.

Kata Kunci: Perikanan Tangkap Laut, Klasterisasi K-Means, Kesenjangan Regional

PENDAHULUAN

Sektor perikanan laut memiliki peran penting dalam mendukung pembangunan ekonomi daerah di Provinsi Jawa Barat, Indonesia. Wilayah pesisir seperti Indramayu, Cirebon, Pangandaran, dan Sukabumi dikenal dengan aktivitas penangkapan ikan yang intensif serta kekayaan sumber daya lautnya. Namun demikian, tingkat produksi aktual masih berada jauh di bawah potensi yang diestimasikan, dan terdapat disparitas produktivitas yang signifikan antar kabupaten/kota pesisir. Ketimpangan ini mencerminkan belum optimalnya pemanfaatan sumber daya laut serta kurangnya kebijakan spesifik wilayah yang mampu menjawab tantangan produksi secara lokal (R. Bussalam, 2024).

Berdasarkan data perikanan nasional, potensi tangkapan gabungan di Wilayah Pengelolaan Perikanan (WPP) 573 dan 712 di Jawa Barat mencapai lebih dari dua juta ton per tahun. Meski demikian, banyak daerah yang belum mampu mencapai performa optimal akibat perbedaan dalam volume hasil tangkapan, jumlah kapal, jumlah nelayan, harga jual ikan, serta jumlah rumah tangga perikanan (D. A. Kurniawan, 2024), (BPS, 2023). Wilayah yang memiliki infrastruktur terbatas atau nilai pasar yang rendah cenderung mengalami kesulitan dalam meningkatkan produksi dan daya saingnya..

Penelitian sebelumnya telah menggunakan teknik data mining seperti K-Means Clustering untuk mengklasifikasikan wilayah produksi perikanan berdasarkan karakteristik yang serupa. Metode pembelajaran tak terawasi ini terbukti efektif dalam mengidentifikasi kelompok produksi yang berbeda serta memberikan dasar bagi penyusunan rencana pengembangan strategis (al, 2023), (S. Gymnastiar, 2024). Namun, sebagian besar penelitian yang ada masih berfokus pada skala nasional atau provinsi, tanpa membahas secara rinci disparitas di tingkat sub-provinsi, khususnya antara wilayah pesisir utara dan selatan Jawa Barat.

Penelitian ini menggunakan algoritma K-Means Clustering untuk mengelompokkan kabupaten dan kota pesisir di Jawa Barat berdasarkan lima variabel utama, yaitu volume produksi ikan, jumlah kapal, jumlah nelayan laut, harga rata-rata per kilogram, dan jumlah rumah tangga perikanan. Tujuannya adalah membangun model segmentasi berbasis data yang mendukung penyusunan rekomendasi kebijakan secara terarah untuk meningkatkan produktivitas perikanan laut.

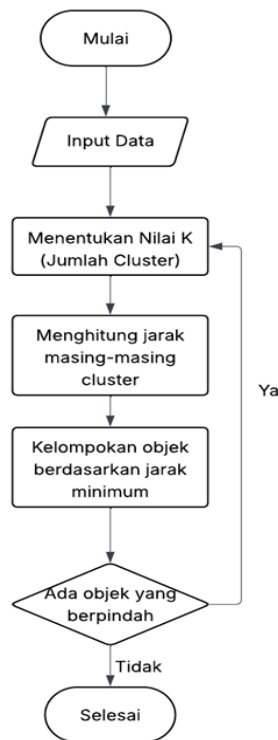
METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan menerapkan teknik pengelompokan data (clustering) (M. Matdoan, 2023) untuk menganalisis karakteristik produksi perikanan laut di kabupaten dan kota pesisir Provinsi Jawa Barat. Data yang digunakan dalam penelitian ini mencakup lima variabel, yaitu volume produksi ikan (ton), jumlah kapal laut (unit), jumlah nelayan laut (orang), harga rata-rata ikan (Rp/kg), dan jumlah rumah tangga perikanan. Kelima variabel tersebut dipilih karena dianggap mewakili kapasitas produksi serta nilai ekonomi dari sektor perikanan. Sebelum dilakukan analisis, data melalui proses pembersihan (data cleaning) untuk menangani nilai yang hilang (missing values) serta memastikan konsistensi antar variabel.

Pengumpulan data merupakan langkah awal untuk memahami suatu fenomena atau permasalahan dengan mengekstraksi informasi yang relevan dari berbagai sumber (Kumari, 2022). Dalam penelitian ini, data yang digunakan merupakan data sekunder yang diperoleh dari situs resmi Kementerian Kelautan dan Perikanan. Dataset yang digunakan terdiri atas 19 kolom dan 31 baris, yang mencakup berbagai variabel yang digunakan dalam proses clustering.

Data yang telah dikumpulkan dianalisis menggunakan algoritma K-Means Clustering. Analisis ini mencakup penentuan jumlah kluster (nilai K), perhitungan jarak antara setiap titik data ke pusat kluster (centroid) dengan rumus jarak seperti Euclidean distance, dan pengelompokan data berdasarkan jarak minimum. Proses ini dilakukan secara iteratif hingga kluster mencapai kondisi stabil (tidak ada lagi data yang berpindah kluster) (V. Winland, 2024).

$$d_{euclidean}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$



Gambar 1. Flowchart Analisis K-Mean Clustering

1. Input Data

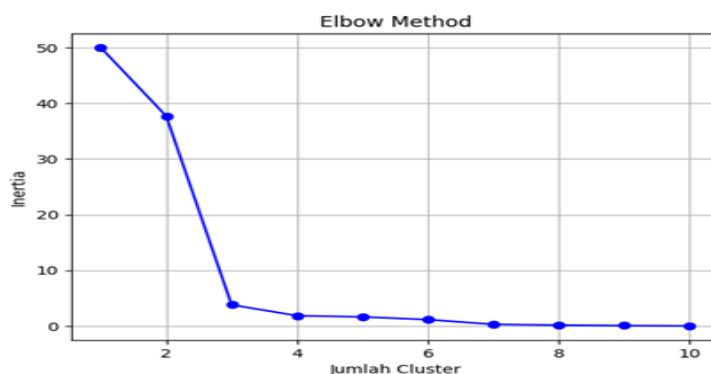
Provinsi	Kab / Kota	Kelompok Ikan	Volume (ton)	Harga Rata-Rata Tertimbang (Rp/kg)	Tahun	Kelompok Ikan.1	Volume (ton).1	Harga Rata-Rata Tertimbang (Rp/kg).1	Tahun.1	...	jumlah kapal	jenis	tahun	Jumlah nelayan laut	tahun.1	Jumlah produksi ikan (ton)
Jawa Barat	Cianjur	Tongkol	10.69	20000	2023	Tenggiri	1.74	40201	2023	...	962	laut	2023	1119	2023	16.29
Jawa Barat	Garut	Tongkol	908.27	15000	2023	Tenggiri	195.83	48793	2023	...	1062	laut	2023	4833	2023	1104.10
Jawa Barat	Indramayu	Tongkol	33804.70	20861	2023	Tenggiri	34881.54	48239	2023	...	4583	laut	2023	41656	2023	83221.03
Jawa Barat	Karawang	Tongkol	360.53	23000	2023	Tenggiri	48.81	51318	2023	...	1332	laut	2023	500	2023	1218.62
Jawa Barat	Kota Cirebon	Tongkol	7.74	29835	2023	Tenggiri	143.36	48594	2023	...	655	laut	2023	1564	2023	162.30
Jawa Barat	Pangandaran	Tongkol	78.21	18541	2023	Tenggiri	103.17	53430	2023	...	1593	laut	2023	544	2023	181.38

Gambar 2. Dataset

Variabel yang diambil untuk pemodelan clustering k-means yaitu Harga Rata-Rata Tertimbang (Rp/kg), Jumlah produksi ikan, jumlah kapal, Jumlah nelayan laut, Rumah Tangga Perikanan dan Kabupaten / Kota.

2. Menentukan Nilai K (Jumlah Cluster)

Metode Elbow adalah teknik heuristik yang digunakan untuk menentukan jumlah kluster optimal (*optimal k*) dalam analisis clustering, terutama pada algoritma K-Means. Metode ini bekerja dengan memplot nilai Within-Cluster Sum of Squares (WCSS) terhadap berbagai nilai k. Titik pada grafik di mana penurunan WCSS mulai melambat dan membentuk sudut seperti “siku” atau *elbow* dipilih sebagai jumlah kluster yang optimal (H. Hartama, 2024).



Gambar 3. Metode Elbow

3. Menghitung Jarak Masing-Masing Cluster

Random state (atau seed) dalam Sklearn menentukan nilai awal untuk pemilihan centroid secara acak ketika menggunakan metode seperti `init='random'` atau `init='k-means++'`. Dengan menentukan nilai seed yang sama, hasil clustering menjadi reproduksi deterministik dan konsisten di setiap eksekusi (A. Deshpande, 2023).

```
[7] kmeans = KMeans(n_clusters=3, random_state=42, n_init='auto')  
    df['Cluster'] = kmeans.fit_predict(X_scaled)
```

Gambar 4. Random State

Kode K-Means Clustering yang digunakan dalam penelitian ini menetapkan jumlah kluster sebanyak tiga (`n_clusters=3`), berdasarkan hasil dari metode Elbow yang menunjukkan bahwa tiga kluster merupakan jumlah yang optimal untuk mengelompokkan data. Untuk menjaga konsistensi hasil klasterisasi dalam setiap eksekusi kode, nilai `random_state` diset ke 42 (H. Ahmed, 2022). Penentuan nilai ini memastikan bahwa proses inisialisasi centroid awal dilakukan secara deterministik,

sehingga hasil yang diperoleh bersifat reproducible. Selanjutnya, parameter `n_init` diatur ke 'auto', yang memungkinkan library *scikit-learn* (versi 1.2 ke atas) memilih jumlah inisialisasi centroid secara otomatis guna meningkatkan stabilitas hasil. Proses klasterisasi dilakukan melalui fungsi `fit_predict(X_scaled)`, yang mengelompokkan data yang telah diskalakan dan menyimpan hasil klaster ke dalam kolom baru bernama 'Cluster'.



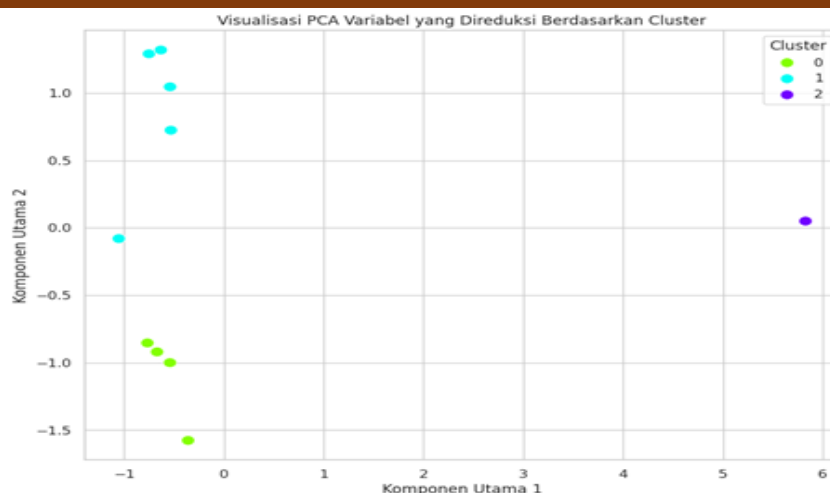
	Kab / Kota	Jumlah produksi ikan (ton)	jumlah kapal	Jumlah nelayan laut	Harga Rata-Rata Gabungan (Rp/kg)	Rumah Tangga Perikanan	Cluster
0	Cianjur	16.29	962	1119	22157.749540	2850	0
1	Garut	1104.10	1062	4833	20993.735341	1089	0
2	Indramayu	83221.03	4583	41656	30641.998917	33347	2
3	Karawang	1218.62	1332	500	15596.623968	5640	0
4	Kota Cirebon	162.30	655	1564	44740.053851	4164	1
5	Pangandaran	181.38	1593	544	38386.066325	1672	1
6	Subang	1194.94	965	3243	44641.254724	2293	1
7	Sukabumi	141.57	6	204	23863.412517	9103	0
8	Tasikmalaya	72.95	253	1670	30812.342975	943	1
9	Bekasi	139.80	1217	4022	41686.000000	1771	1

Gambar 5. Hasil Jarak

Pada gambar 5, Cluster 1 mencakup 5 kabupaten/kota yaitu Kota Cirebon, Pangandaran, Subang, Tasikmalaya dan Bekasi . Kemudian, Cluster 0 mencakup 4 wilayah yaitu Cianjur, Garut, Karawang dan Sukabumi. Cluster 2 mencakup 1 wilayah yaitu Indramayu.

4. Kelompokkan Objek Berdasarkan Jarak Minimum

Visualisasi PCA (Principal Component Analysis) adalah cara untuk merepresentasikan data berdimensi tinggi ke dalam dua atau tiga dimensi utama agar lebih mudah dipahami secara visual. PCA adalah teknik reduksi dimensi yang mengubah variabel-variabel asli menjadi sekumpulan komponen utama yang saling ortogonal (tidak saling berkorelasi), di mana komponen pertama menangkap variabilitas data paling besar, diikuti oleh komponen kedua, dan seterusnya (S. Hamdy Abdel Latif, 2022).



Gambar 6. Visualisasi PCA

Visualisasi PCA pada gambar 6 menunjukkan pemetaan wilayah ke dalam tiga klaster utama. Klaster 2 (ungu) terpisah jauh, menandakan wilayah dengan produksi dan variabel pengaruh. Sementara itu, klaster 0 (hijau) dan klaster 1 (biru) berdekatan, mencerminkan karakteristik produksi dan variabel pengaruh yang lebih rendah dan menengah. Visualisasi ini menegaskan perbedaan tingkat potensi antar wilayah. Metode K-Means memungkinkan identifikasi pola tersembunyi dalam data multivariat, sehingga tiap klaster dapat merepresentasikan segmentasi wilayah dengan pendekatan yang lebih objektif dan berbasis data. Hal ini sangat membantu dalam perumusan kebijakan yang lebih terfokus dan efektif (Ayodele, 2024).

5. Ada Objek Berpindah

Jika dalam hasil ada klaster yang berpindah maka analisis harus dilakukan ulang.

HASIL DAN PEMBAHASAN

Tabel 1. Karakteristik Cluster

Variabel	Cluster 0	Cluster 1	Cluster 2
Jumlah Produksi Ikan (Ton)	16 – 1.219	73 – 1.195	83.221
Jumlah Kapal	6 – 1.332	253 – 1.593	4.583
Jumlah Nelayan Laut	204 – 4.833	544 – 4.022	41.656
Harga Rata-Rata Gabungan (Rp/kg)	15.597 – 23.863	30.812 – 44.740	30.642
Rumah Tangga Perikanan	1.089 – 9.103	943 – 4.164	33.347

Penjelasan:

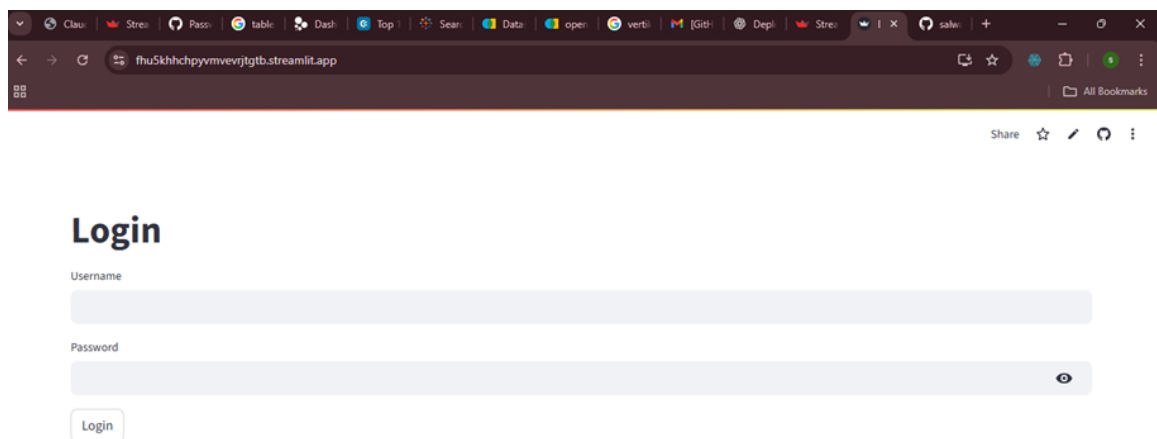
- a. Cluster 0 menunjukkan wilayah-wilayah dengan volume produksi ikan yang relatif sedang. Harga jual ikan di wilayah ini tergolong rendah dibanding klaster lainnya. Aktivitas penangkapan tinggi, tetapi efisiensi perkapal atau nelayan tergolong rendah-menengah. Cluster 0 termasuk dalam kategori “Produksi Aktif Berkinerja Sedang”, yaitu wilayah dengan aktivitas perikanan yang tinggi namun belum optimal dari sisi produktivitas dan nilai ekonomi.
- b. Cluster 1 menunjukkan wilayah dengan harga jual ikan yang tinggi, namun volume produksinya tidak setinggi cluster 0. Jumlah kapal tergolong menengah, nelayan tergolong lebih rendah dan rumah tangga perikanan tergolong rendah dari cluster 0, namun efisiensi hasil tangkap per unit lebih baik. Kondisi ini mengindikasikan wilayah dengan jenis ikan bernilai tinggi dan sistem tangkap yang lebih efisien. Dengan demikian, Cluster 1 termasuk dalam kategori “Efisiensi Tinggi Berbasis Nilai”, yaitu wilayah dengan potensi ekonomi kuat meskipun skala produksi tidak besar.
- c. Cluster ini merupakan outlier karena semua nilai variabelnya sangat tinggi dan tetap, yang menunjukkan hanya satu wilayah dengan dominasi produksi yang sangat besar. Wilayah ini memiliki jumlah kapal, nelayan, rumah tangga perikanan, serta volume produksi yang jauh melampaui wilayah lain. Oleh karena itu, Cluster 2 dikategorikan sebagai “Klaster Dominan Produksi Ekstrem” atau “Outlier Positif”, yaitu wilayah kunci yang secara signifikan berkontribusi terhadap total produksi provinsi dan perlu pengelolaan berkelanjutan.

Tabel 2. Perbandingan Pesisir Selatan dan Utara

Aspek	Zona Selatan	Zona Utara
Wilayah	Cianjur, Garut, Pangandaran, Sukabumi, Tasikmalaya	Karawang, Kota Cirebon, Subang, Bekasi, Indramayu
Cluster	Cluster 0 dan 1	Cluster 0, 1, dan 2
Rata-rata Produksi (ton)	303,26	14.626,04
Rata-rata Jumlah Kapal	775	3.683
Rata-rata Nelayan Laut	1.674	14.707
Rata-rata Rumah Tangga	3.131	11.034
Rata-rata Harga (Rp/kg)	Rp. 27.242	Rp. 34.441

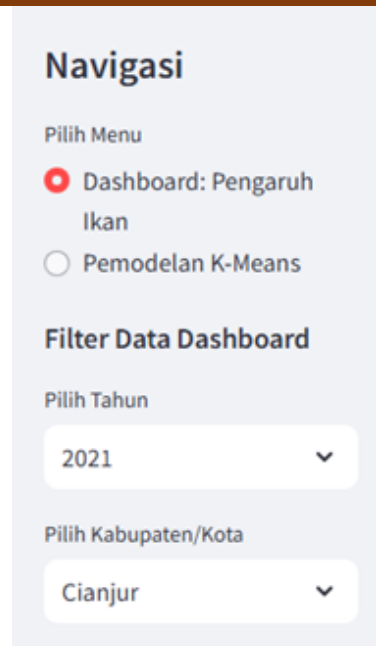
Analisis	Didominasi oleh perikanan tangkap skala kecil hingga menengah. Sarana tangkap dan pengolahan belum merata.	Didominasi oleh perikanan skala besar dan terorganisasi. Telah berkembang hingga sektor hilir dan distribusi.
Rekomendasi	Perlu dukungan dasar seperti alat tangkap, pelatihan, dan penguatan komunitas.	Fokus pada efisiensi industri, pengolahan, distribusi, dan perlindungan nelayan.

Proses analisis data dan perumusan rekomendasi diselesaikan sebagai bagian akhir dari prosedur penelitian dengan dashboard analisis agar pengguna bisa akses untuk rekomendasi di masing-masing wilayah.

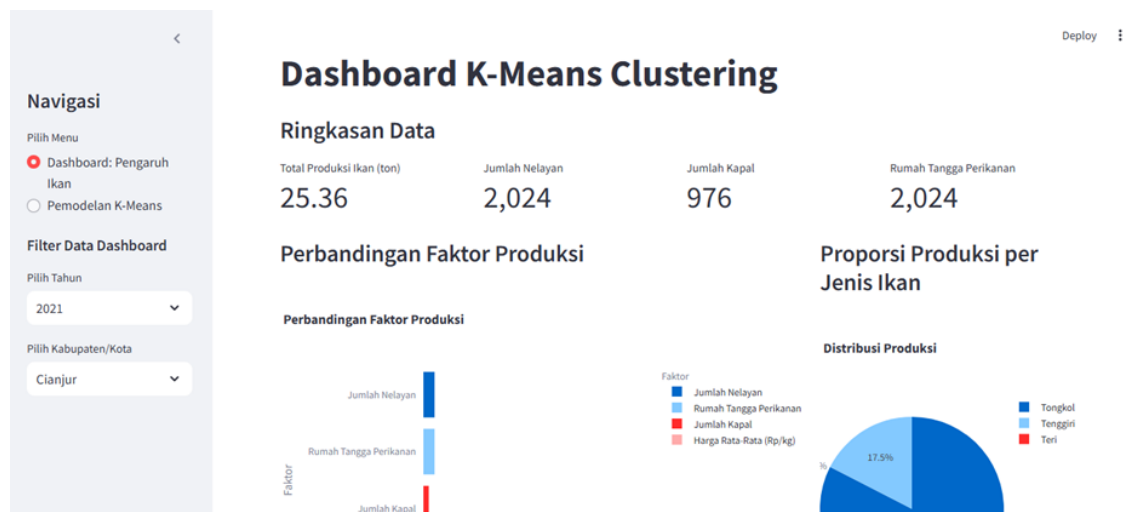


Gambar 7. Tampilan Awal Dashboard

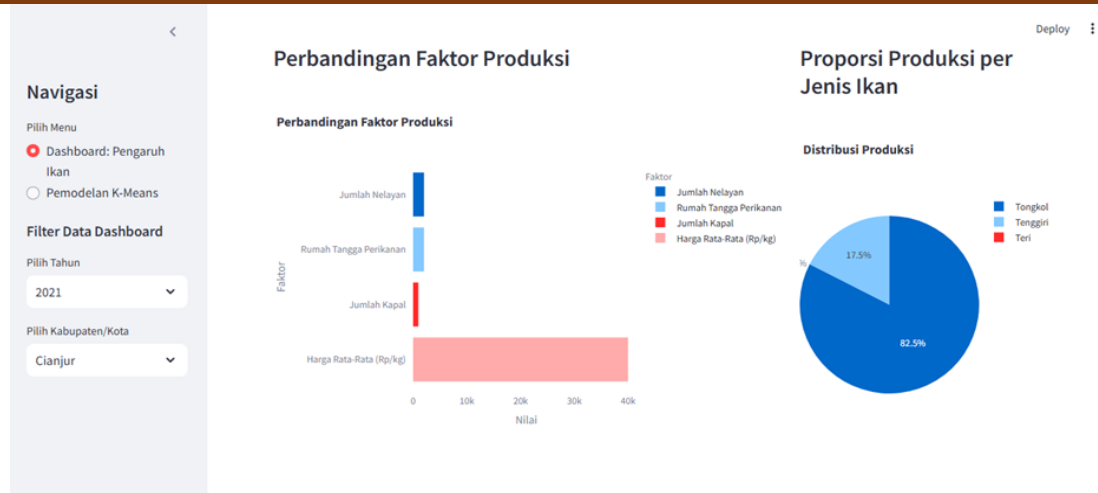
Pada tampilan awal dashboard, pengguna diminta untuk melakukan login terlebih dahulu. Sebagai contoh, pada halaman login ini dibuat akun dengan username: admin dan password: admin untuk dapat masuk ke dalam sistem.



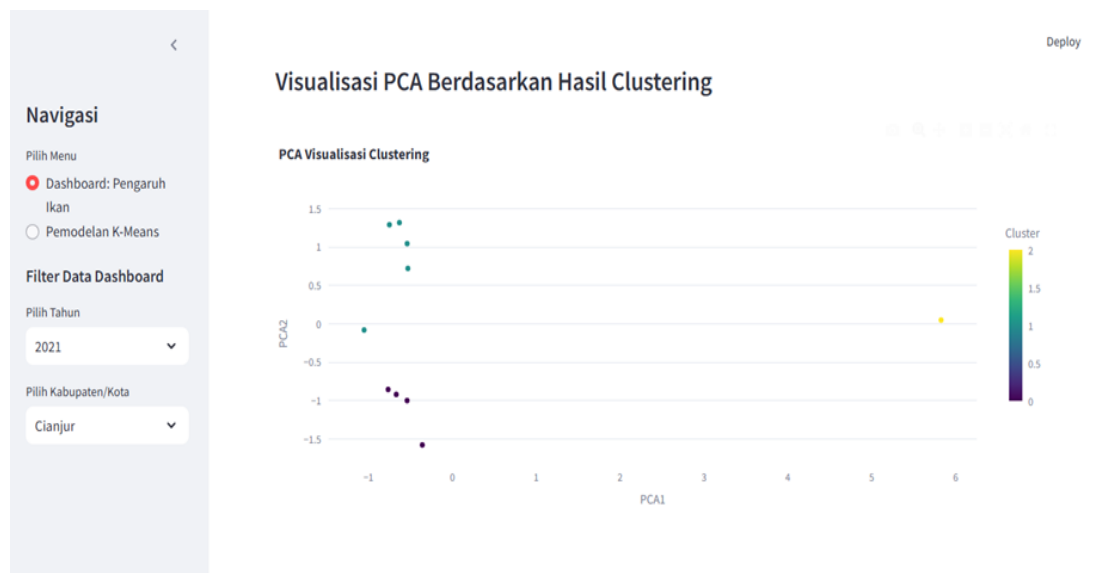
Gambar 8. Fitur Navigasi



Gambar 9. Tampilan Menu Dashboard



Gambar 10. Tampilan Menu Dashboard



Gambar 11. Tampilan Menu Dashboard

Pada menu Dashboard di gambar 10 dan 11, ditampilkan berbagai visualisasi dari variabel-variabel dalam data, seperti big number, bar chart, dan pie chart. Visualisasi seperti big number, diagram batang (bar chart), dan diagram lingkaran (pie chart) digunakan dalam dashboard untuk menyajikan variabel-variabel penting secara jelas dan intuitif (W. Perkins, 2023). Big number menampilkan metrik agregat secara langsung, bar chart memudahkan perbandingan antarwilayah, dan pie chart memperlihatkan proporsi kontribusi masing-masing wilayah. Visualisasi ini dirancang agar pengguna dapat memahami informasi dengan

cepat dan mendukung proses pengambilan keputusan yang lebih tepat (Y. LI, 2024). Pengguna dapat memanfaatkan fitur filter berdasarkan tahun dan wilayah untuk menyesuaikan tampilan visualisasi tersebut sesuai kebutuhan analisis. Namun, filter tahun dan wilayah tidak berlaku untuk visualisasi K-Means Clustering. Hal ini dikarenakan hasil clustering didasarkan pada rata-rata nilai variabel selama tahun 2021 hingga 2023, sehingga bersifat agregat dan tidak terpengaruh oleh filter waktu maupun wilayah.

Deploy

⋮

Statistik Rata-rata per Cluster

Cluster	Jumlah produksi ikan [ton]			Jumlah kapal			Jumlah nelayan laut			Harga Rata-Rata Gabungan [Rp/kg]			Rumah Tangga Perikanan		
	mean	min	max	mean	min	max	mean	min	max	mean	min	max	mean	min	max
0	620.14	16.29	1,218.62	840.5	6	1,332	1,664	204	4,833	20,652.88	15,596.62	23,863.41	4,670.5	1,089	9,164
1	350.27	72.95	1,194.94	936.6	253	1,593	2,208.6	544	4,022	40,053.14	30,812.34	44,740.05	2,168.6	943	4,164
2	83,221.03	83,221.03	83,221.03	4,583	4,583	4,583	41,656	41,656	41,656	30,641.99	30,641.99	30,641.99	33,347	33,347	33,347

Wilayah & Cluster

	Kab / Kota	Jumlah produksi ikan (ton)	jumlah kapal	Jumlah nelayan laut	Harga Rata-Rata Gabungan (Rp/kg)	Rumah Tangga Perikanan
0	Cianjur	16.29	962	1,119	22,157.75	2,850
1	Garut	1,104.1	1,062	4,833	20,993.74	1,089
2	Indramayu	83,221.03	4,583	41,656	30,641.99	33,347
3	Karawang	1,218.62	1,332	500	15,596.62	5,640
4	Kota Cirebon	162.3	655	1,564	44,740.05	4,164
5	Pangandaran	181.38	1,593	544	38,386.07	1,672

Gambar 12. Tampilan Menu Pemodelan K-Means

Deploy

⋮

Filter Rekomendasi Per Wilayah

Pilih Wilayah:
Cianjur

Wilayah: Cianjur

Masuk Cluster: 0

	Jumlah produksi ikan (ton)	jumlah kapal	Jumlah nelayan laut	Harga Rata-Rata Gabungan (Rp/kg)	Rumah Tangga Perikanan
0	16.29	962	1,119	22,157.75	2,850

Rekomendasi:

- Rekomendasi: peningkatan jumlah kapal & akses pasar.

Gambar 13. Tampilan Menu Dashboard

Pada menu Pemodelan K-Means di gambar 12 dan 13, ditampilkan hasil pemodelan clustering terhadap data produksi ikan. Menu ini mencakup tampilan dataset, hasil klasterisasi pada masing-masing wilayah, serta rata-rata nilai variabel pada setiap klaster. Selain itu, tersedia juga fitur filtering rekomendasi yang memberikan saran peningkatan produksi ikan berdasarkan klaster yang terbentuk. Dengan demikian, pengguna dapat melihat karakteristik wilayah berdasarkan klaster dan menentukan strategi yang sesuai untuk peningkatan produksi.

KESIMPULAN DAN REKOMENDASI

Berdasarkan hasil analisis K-Means terhadap data produksi ikan laut di kabupaten/kota pesisir Jawa Barat, wilayah-wilayah dikelompokkan menjadi tiga klaster yang merepresentasikan karakteristik berbeda dengan variabel Produksi Ikan, Jumlah Nelayan, Jumlah Kapal Laut, Harga Rata-Rata (Kg/Rp) dan Rumah Tangga Perikanan.

Cluster 0 mencakup wilayah-wilayah Cianjur, Garut, Karawang, dan Kota Cirebon, menunjukkan aktivitas perikanan yang cukup tinggi namun belum disertai dengan efisiensi penangkapan yang optimal. Oleh karena itu, diperlukan pelatihan peningkatan keterampilan nelayan, modernisasi alat tangkap, serta penguatan akses pasar dan manajemen usaha perikanan agar produktivitas dan nilai ekonomi wilayah ini dapat ditingkatkan secara berkelanjutan.

Cluster 1 mencakup wilayah-wilayah Kabupaten Pangandaran, Subang, Sukabumi, dan Bekasi, menunjukkan karakteristik efisiensi penangkapan ikan yang relatif tinggi serta nilai jual hasil tangkapan yang unggul. Oleh karena itu, strategi pengembangan yang tepat untuk klaster ini adalah dengan memperkuat komoditas unggulan bernilai tinggi, melakukan diversifikasi produk hasil perikanan, serta mendorong pemanfaatan teknologi pascapanen dan perluasan akses pasar.

Cluster 2 hanya terdiri dari satu wilayah yaitu Indramayu yang menjadi outlier positif dengan dominasi produksi yang sangat besar dibanding wilayah lain. Oleh karena itu, wilayah ini membutuhkan kebijakan yang berorientasi pada pengelolaan perikanan yang berkelanjutan, termasuk pengendalian kapasitas tangkap, investasi pada infrastruktur rantai dingin, serta efisiensi logistik distribusi hasil laut agar daya saing dan keberlanjutan sektor

REFERENSI

- Ayodele, E. (2024). Through unsupervised learning, K means reveals hidden patterns in customer data, enabling personalized communications. *Int. J. Creative Research Thoughts*, 12(7).
- Abadi, M. (2016). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*. Retrieved from <http://arxiv.org/abs/1603.04467>
- al, A. M. (2023). K-Means Clustering Algorithm: A Comprehensive Review, Variants, and Developments in the Big Data Era. *Information Sciences*, 662, 178-210.
- A. Deshpande, R. P. (2023, Sep). *Improved Outlier Robust Seeding for k-means*. Retrieved from arXiv preprint arXiv.
- BPS. (2023). *Statistical Terminology – 2023 Agricultural Census: Marine Fisheries Households*. Retrieved from Badan Pusat Statistik: <https://sensus.bps.go.id>
- Collobert, R., & Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. *Proc. 25th Int. Conf. Mach. Learn.*, (pp. 160-167).
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018, October). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. Retrieved from <http://arxiv.org/abs/1810.04805>
- D. A. Kurniawan, E. S. (2024). Mapping competitiveness of the marine capture fishery sub-sector in West Java Province. *Marine and Aquaculture Research Media*, 19(1), 1-12.
- Graves, A. (2013). *Generating Sequences With Recurrent Neural Networks*. Retrieved from <http://arxiv.org/abs/1308.0850>
- H. Hartama, T. S. (2024). Clustering Wilayah Rawan Bencana Alam Menggunakan K-Means dan K-Medoids dengan Evaluasi Elbow, DBI dan Silhouette. *Jurnal Bina Informatika*, 11(1), 15-22.
- H. Ahmed, M. D. (2022). Managing Randomness to Enable Reproducible Machine Learning Experiments. *International Journal of Experimental Machine Learning*, 14(3), 45-60.
- Jurafsky, J. H., & Daniel. (2017). *Speech and Language Processing*. Ghost Danc. doi:10.2307/j.ctt1djmg7w.13.

- Kim, Y. (2014). Convolutional neural networks for sentence classification. *Conf. Empir. Methods Nat. Lang. Process. Proc. Conc* (pp. 1746-1751). EMNLP. doi:10.3115/v1/d14-1181
- Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.* ICLR.
- Kumari, A. (2022). A Review of Research Process: Data Collection and Analysis. *Galaxy International Interdisciplinary Research Journal*, 10(5), 771-780.
- Liu, B., & Lane, I. (2016). Attention-based recurrent neural network models for joint intent detection and slot filling. *Proc. Annu. Conf. Int. Speech Commun. Assoc.* (pp. 685-689). INTERSPEECH. doi:10.21437/Interspeech.2016-135
- M. Matdoan, M. W. (2023). Application of the K-Means Algorithm for Clustering Production of Capture Fisheries in Maluku Province. *Pattimura Journal of Mathematics*, 12(1), 41-50.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013, January). *Efficient Estimation of Word Representations in Vector Space*. Retrieved from <http://arxiv.org/abs/1301.3781>
- N., A. Y., C. P. Richard Socher, A. P., & Manning, C. D. (2004). Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank Richard. *Empir. Methods Nat. Lang. Process.*, 1631-1642.
- R. Bussalam, J. (2024). Analysis of the distribution of marine capture fishery production potential in West Java Province. *Journal of Marine and Fisheries of Sulawesi*, 8(1), 1-10.
- Sahu, A., Pandey, S., Aarwal, M., & Chauhan, S. (2023). Offline Virtual Chat Bot by Using Natural Language Processing. *SSRN Electron.* doi:10.2139/ssrn.4386503
- S. Hamdy Abdel Latif, A. S. (2022). Principal component analysis as tool for data reduction with an application. *EUREKA: Physics and Engineering*(5), 184-198.
- V. Winland, E. K. (2024, Jun). *What is K-Means Clustering?* Retrieved from IBM Think.
- W. Perkins, V. B. (2023, Juli). *The Power of Big Aggregate Numbers*. Retrieved from DataVis Community Blog.
- Y. B, I. G., & Aaron, C. (2016). *Deep Learning*. Retrieved from www.deeplearningbook.org

Y. LI, E. B. (2024, okt). *Does the Infamous Pie Chart Really Hurt Decision-Making in the Real World? Assessing the Role of Visualization in High Level Academic Decisions*. Retrieved from arXiv Preprint.