

Clustering Data Demografi Penduduk Daerah Pesisir Sulawesi Selatan Menggunakan Metode K-Means

Vitha Aulia Nurwardhani Yunus, Keziah Julia Siow, Rismayani, Sriwahyuningsih Piu*

Jurusan Teknik Informatika, Universitas DIPA Makassar, Makassar

Jl. Perintis Kemerdekaan KM.9 Tamalanrea, Kota Makassar, (0411) 587194

e-mail: ¹vitha.yunus@gmail.com, ²keziahjuliasow@gmail.com, ³rismayani@undipa.ac.id, ⁴sri.wahyuningsih@undipa.ac.id.

Abstrak

Wilayah pesisir Sulawesi Selatan memiliki banyak potensi ekonomi biru, terutama dalam bidang perikanan dan pariwisata. Namun, kurangnya pemahaman tentang karakteristik demografi dan potensi masing-masing wilayah menyebabkan pemanfaatan potensi ini tidak optimal. Penelitian ini menggunakan metode K-Means untuk membuat strategi pengembangan ekonomi biru yang tepat sasaran dengan mengelompokkan wilayah pesisir di Sulawesi Selatan berdasarkan karakteristik demografi. Data sekunder sebanyak 501 data desa/kelurahan dianalisis menggunakan Google Colab dan Python. Proses preprocessing data dilakukan melalui penghapusan duplikat, penanganan nilai kosong, deteksi dan penghapusan outlier, serta normalisasi. Dengan menggunakan metode elbow, ditemukan kluster terbaik dengan hasil 2 (dua) kluster berdasarkan jumlah penduduk, tingkat pendidikan, usia produktif, dan sektor pekerjaan (perdagangan, nelayan, dan wiraswasta). Dengan nilai silhouette score 0,33 yang menunjukkan tingkat separasi antar kluster yang relatif rendah. Nilai ini mengindikasikan bahwa sebagian wilayah memiliki karakteristik demografi yang saling tumpang tindih, sehingga batas antar kluster tidak sepenuhnya tegas.

Kata kunci—Ekonomi Biru, Algoritma K-Means, Silhouette Score, Demografi, Sulawesi Selatan.

Abstract

The coastal areas of South Sulawesi have a lot of blue economic potential, especially in fisheries and tourism. However, a lack of understanding of the demographic characteristics and potential of each region leads to suboptimal utilization of this potential. This study uses the K-Means method to create a targeted blue economy development strategy by clustering coastal areas in South Sulawesi based on demographic characteristics. Secondary data of 501 villages were analyzed using Google Colab and Python. Data preprocessing was done through duplicate removal, empty value handling, outlier detection and removal, and normalization. By using the elbow method, the best cluster was found with the results of 2 (two) clusters based on population, education level, productive age, and employment sector (trade, fishing, and self-employed). With a silhouette score of 0.33, which indicates a relatively low level of separation between clusters.

Keywords—Blue Economy, K-Means Algorithm, Silhouette Score, Demographics, South Sulawesi complexity.

1. PENDAHULUAN

Daerah pesisir Sulawesi Selatan merupakan wilayah yang kaya akan sumber daya alam dan memiliki potensi ekonomi yang signifikan, terutama dalam sektor perikanan dan pariwisata. Dengan garis pantai yang panjang dan beragam ekosistem laut, masyarakat pesisir di Sulawesi

Selatan memiliki ketergantungan yang tinggi terhadap sumber daya laut untuk mata pencaharian. Dalam konteks ini, konsep ekonomi biru atau blue economy menjadi sangat relevan, karena berfokus pada pemanfaatan sumber daya laut secara berkelanjutan untuk meningkatkan kesejahteraan masyarakat pesisir [1]. Ekonomi biru tidak hanya mencakup pengelolaan sumber daya perikanan, tetapi juga melibatkan pengembangan sektor-sektor lain seperti pariwisata ekologi yang semuanya dapat memberikan manfaat ekonomi yang lebih luas bagi masyarakat pesisir.

Masalah dalam penelitian ini adalah belum adanya informasi atau analisis yang menjelaskan potensi desa, terutama dalam aspek demografi dan sektor ekonomi seperti jumlah penduduk, usia produktif, tingkat pendidikan, dan jenis mata pencarian, serta apakah data tersebut tergolong tinggi atau pun rendah berdasarkan hasil pendataan. Meskipun potensi yang besar, masyarakat pesisir Sulawesi Selatan menghadapi berbagai tantangan modern yang mengancam keberlanjutan dan potensi manfaatnya [2]. Pendataan potensi desa dilaksanakan oleh Badan Pusat Statistik Provinsi Sulawesi Selatan setiap 3 kali dalam 10 tahun [3]. Akan tetapi, dari hasil pendataan ini tidak dijelaskan mengenai wilayah yang memiliki potensi desa yang tinggi maupun rendah.

Untuk mengatasi masalah belum adanya informasi mengenai wilayah yang memiliki potensi desa tinggi maupun rendah, penelitian ini bertujuan mengidentifikasi keadaan potensi desa dengan menggunakan metode analisis kluster untuk mengelompokkan data demografi penduduk berdasarkan karakteristik yang relevan. Salah satu solusi yang dapat diterapkan adalah penggunaan metode K-Means dalam clustering data demografi penduduk daerah pesisir Sulawesi Selatan. Dengan mengidentifikasi kelompok-kelompok penduduk yang memiliki potensi desa serupa, dapat digunakan untuk merancang kebijakan ekonomi biru yang lebih tepat sasaran dan responsif terhadap kondisi lokal dan menjaga keberlanjutan sumber daya laut yang mereka andalkan. Dengan demikian, integrasi antara analisis demografi dan prinsip-prinsip ekonomi biru dapat menciptakan sinergi yang bermanfaat bagi masyarakat pesisir di Sulawesi Selatan.

Adapun beberapa pustaka yang menjadi rujukan dalam penelitian ini adalah:

Ekonomi biru merupakan ekonomi laut bertujuan untuk meningkatkan kesejahteraan manusia dan kesetaraan manusia dan kesetaraan sosial, sambil secara signifikan mengurangi risiko lingkungan dan kelangkaan ekologi. Bank dunia mendefinisikan “Ekonomi Biru” sebagai penggunaan sumber daya laut secara berkelanjutan untuk pertumbuhan ekonomi, peningkatan mata pencaharian dan lapangan kerja sambil menjaga kesehatan ekosistem laut [4].

Data pada dasarnya adalah kumpulan statistik dan fakta yang masih mentah yang perlu dianalisis untuk memiliki makna atau digunakan sebagai referensi. Dengan demikian, data berfungsi sebagai jembatan antara proses pengolahan informasi di dunia digital dengan peristiwa di dunia nyata [5], [6].

Clustering merupakan metode yang bersifat unsupervised atau tanpa arahan di mana karakteristik tiap cluster tidak ditentukan sebelumnya, melainkan menurut kemiripan atribut-atribut dari suatu kelompok. K-Means merupakan salah satu algoritma clustering yang digunakan dalam proses Data mining. K-Means bekerja dengan cara melakukan pengelompokan secara partisi yang memisahkan data ke dalam kelompok-kelompok tertentu dengan meminimalkan rata-rata jarak setiap data klasternya [7]

Python merupakan bahasa pemrograman yang berfokus pada keterbacaan kode, menggabungkan kemampuan, kapabilitas, memiliki kode pemrograman yang jelas serta dilengkapi oleh banyak fungsi library yang banyak dan komprehensif. Python juga mendukung multi-paradigma program yang secara umum berorientasi pada pemrograman fungsional, pemrograman berbasis objek dan pemrograman imperatif [8]

Google Colab atau Google Collaboratory merupakan tools atau code editor yang menggunakan sistem penyimpanan awan yang gratis untuk tujuan penelitian. Google Colab dirancang dengan elemen dari Jupyter, dan semua masukan library hampir semua tersedia untuk digunakan dalam sebuah penelitian termasuk meneliti menggunakan data mining. Berdasarkan kesamaan dan kegunaan fungsi, Google Colab sama seperti Jupyter Notebook dikarenakan kegunaan dan fungsinya yang sama, namun perbedaannya terdapat pada pembatasan sistem karena Google Colab dapat dijalankan secara gratis dan online [9].

Adapun beberapa penelitian yang terkait yaitu:

Penelitian yang dilakukan oleh Syaputri dkk dengan judul Implementasi Algoritma K-Means untuk pengelompokan distribusi sosial ekonomi masyarakat berdasarkan demografi kependudukan, yaitu menggunakan data kuesioner dari perumahan tertentu, yang mencakup pendapatan, pendidikan, dan jumlah tanggungan. Hasil analisis menunjukkan pola hubungan antara karakteristik demografi dan kondisi sosial ekonomi [10].

Kemudian penelitian yang dilakukan oleh Azrahwati dkk dengan judul K-Means cluster analysis for grouping districts in south sulawesi province based on village potential, yaitu penerapan K-Means untuk mengelompokkan kecamatan di Provinsi Sulawesi Selatan berdasarkan potensi desa [11].

Selanjutnya penelitian yang dilakukan Cisneros-Montemayor dkk dengan judul enabling conditions for an equitable and sustainable blue economy, yaitu eksplorasi kondisi yang diperlukan untuk membangun ekonomi biru yang berkelanjutan dan adil [1].

Selanjutnya penelitian yang dilakukan oleh Lee dkk dengan judul penelitian the blue economy and the united nations' sustainable development goals: challenges and opportunities, yaitu mengkaji hubungan antara ekonomi biru dan Tujuan Pembangunan Berkelanjutan [4].

Selanjutnya penelitian yang dilakukan oleh Darmin dengan judul pendekatan CEPA (Collaborative, Ecology, Planning, Administrative) dalam pemberdayaan masyarakat pesisir pantai, yaitu membahas mengenai peningkatan akses dan kualitas pendidikan di daerah pesisir dapat berkontribusi pada peningkatan kesejahteraan masyarakat [12].

Berbeda dari penelitian-penelitian sebelumnya yang lebih menitikberatkan pada kondisi umum ekonomi biru, hubungan dengan TPB, atau pendekatan pemberdayaan masyarakat, penelitian ini memperkenalkan inovasi dengan mengelompokkan wilayah berdasarkan potensi desa melalui metode K-Means.

2. METODE PENELITIAN

Metode penelitian diterapkan untuk menyediakan dasar yang sistematis dalam mencapai tujuan penelitian. Dengan menggunakan pendekatan yang sistematis, penelitian ini berusaha mendapatkan hasil yang valid dan dapat dipertanggungjawabkan. Adapun metode dari penelitian ini adalah:

2.1 Waktu dan Tempat Penelitian

Penelitian ini berlangsung dari Oktober 2024 hingga Januari 2025. Karena data dapat diakses dari internet, penelitian dapat dilakukan dari mana saja baik secara *online* maupun *offline*.

2.2 Jenis Penelitian

Penyelesaian penelitian ini menggunakan metode penelitian kuantitatif menggunakan data sekunder. Metode penelitian kuantitatif dengan data sekunder adalah pendekatan yang memanfaatkan data yang telah dikumpulkan sebelumnya oleh pihak lain, seperti lembaga pemerintah, organisasi, atau penelitian terdahulu. Dalam metode ini, penulis tidak melakukan pengumpulan data langsung, melainkan menganalisis data yang tersedia untuk menjawab pertanyaan penelitian.

2.3 Sumber Data

Penelitian dilakukan dengan menggunakan data sekunder, yaitu data yang sudah ada sebelumnya dan dikumpulkan, serta sudah tersedia di situs resmi Direktorat Jenderal Kependudukan dan Pencatatan Sipil – Kementerian Dalam Negeri. *Website* tersebut dapat dilihat dalam URL “<https://gis.dukcapil.kemendagri.go.id/>”. Secara spesifik, data yang digunakan meliputi data kependudukan tingkat desa di wilayah pesisir Sulawesi Selatan.

2.4 Metode Pengumpulan Data

Pendekatan pengumpulan data dalam penelitian ini adalah data spasial sekunder, yang memanfaatkan data demografi yang sudah ada sebelumnya seperti peta digital, data statistik, atau

basis data dari lembaga pemerintah atau organisasi lainnya. Data spasial yang diperoleh dari peta digital disajikan dalam bentuk tabel untuk setiap unit administratif kelurahan/desa.

Keberhasilan penerapan strategi ekonomi biru yang berkelanjutan bergantung pada komponen seperti usia, pendidikan, mata pencaharian masyarakat pesisir (nelayan, perdagangan, dan wiraswasta), dan perpindahan penduduk [13], [14], [15], maka dari *dataset* yang ada, dipilih jumlah penduduk, perpindahan penduduk, usia produktif (15-64), dan pendidikan (SLTA ke Atas) sebagai kriteria yang digunakan dalam *clustering*, dengan 501 baris data.

2.5 Metode Analisis

Analisis dilakukan dengan menerapkan metode K-Means untuk mengelompokkan desa pesisir berdasarkan data demografi kependudukan di Sulawesi Selatan. K-Means. Algoritma *clustering K-Means* dapat membagi data berdasarkan jarak antar data pada kelompok yang telah ditetapkan. Penetapan kelompok dalam satu *cluster* dapat dilakukan dengan menghitung jarak setiap objek dengan titik pusat. Algoritma ini bergantung pada fungsi untuk mengukur data yang mempunyai ciri khas sama. Jarak itu sendiri dihitung menggunakan fungsi *Euclidean*. Kemudian data dimasukkan dalam kelompok yang mempunyai jarak terdekat [16].

2.6 Metode Pengujian

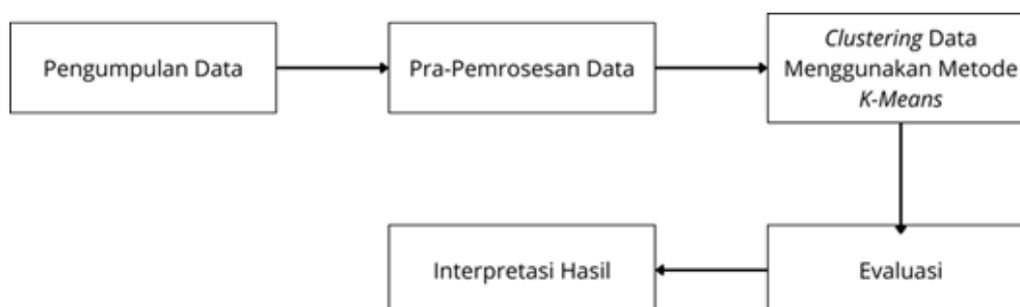
Penelitian ini menggunakan silhouette score sebagai metode pengujian untuk menunjukkan seberapa dekat data dalam satu klaster dan seberapa jauh jarak mereka satu sama lain. Silhouette Score menghitung semua rata-rata silhouette coefficient dalam *dataset*. Rata-rata silhouette coefficient yang tinggi menunjukkan bahwa ada pengelompokan yang baik.

3. HASIL DAN PEMBAHASAN

3.1 Validasi Data

Berdasarkan koordinasi dengan pihak Badan Pusat Statistik (BPS) Provinsi Sulawesi Selatan, mengkonfirmasi bahwa data yang dibutuhkan untuk penelitian ini tersedia dalam publikasi resmi mereka. Namun, data yang digunakan dalam penelitian ini diambil dari DUKCAPIL melalui situs web mereka di <https://gis.dukcapil.kemendagri.go.id/peta/>, untuk detail lebih rinci tentang kelurahan atau desa. Dalam publikasi teknis BPS menjelaskan proses sinkronisasi data antara DUKCAPIL dan BPS dalam upaya untuk mencapai Satu Data Kependudukan, sehingga penggunaan kedua sumber ini dapat diandalkan dalam analisis pengelompokan data demografi penduduk Sulawesi Selatan.

3.2 Perancangan Solusi



Gambar 1 Perancangan Solusi

Adapun penjelasan untuk simbol alur perancangan solusi sebagai berikut:

3.2.1 Pengumpulan Data

Data yang diambil dari *website* kemudian dikumpul dalam 1 *dataset spreadsheet*, seperti pada Tabel 1.

Tabel 1 Dataset Awal

No.	Provinsi	Kabupaten /Kota	Kecamatan	Kelurahan /Desa	Jumlah Penduduk	Perpindahan Penduduk	Usia Produktif (15-64)	..	Wiraswasta
1	Sulawesi Selatan	Kab. Jeneonto	Tarawang	Tarowang	2280	86	1555	..	130
2	Sulawesi Selatan	Kab. Jeneonto	Tarawang	Balangloe Tarowang	3381	16	2281	..	150
...
501	Sulawesi Selatan	Kab. Luwu Utara	Tana Lili	Munte	2197	15	1432	..	137

3.2.2 Pra Pemrosesan Data

Adapun beberapa pra-pemrosesan data yang dilakukan sebagai berikut:

a. Data duplikat

Tahap pra-pemrosesan data untuk menangani data duplikat menggunakan *tools* Google Colab. Setelah mendeteksi data duplikat, terdapat 6 data duplikat yakni, Desa Sambali, Ujung Labuang, Pabbentengang, Mangempang, dan Carigading. Data duplikat ini kemudian dihapus menggunakan fungsi *drop_duplicates()*, fungsi ini akan menghapus baris yang memiliki nilai yang sama persis di semua kolom, kecuali untuk baris pertama yang muncul. Baris pertama yang muncul akan dipertahankan, sedangkan baris-baris duplikat berikutnya akan dihapus menyisakan data berjumlah 495.

b. Nilai Kosong

Pada tahap ini terdapat data dengan nilai kosong (*missing value*) pada atribut nelayan, wiraswasta, dan perdagangan. *Missing value* bisa diatasi dengan pengisian *mean* pada data.

c. Outlier

Outlier dapat didefinisikan sebagai amatan yang menyimpang sedemikian jauh pengamatan lainnya. Hasil deteksi *outlier* menunjukkan bahwa terdapat 24 *outlier* pada variabel Jumlah Penduduk dan Usia Produktif (15-64), 30 *outlier* pada variabel Nelayan dan Pendidikan (SLTA ke Atas), 29 *outlier* pada variabel Perdagangan, Pendidikan (SMA ke atas), 32 *outlier* pada Perpindahan Penduduk. Variabel wiraswasta memiliki *outlier* paling banyak, yaitu 43 data. Setelah penghapusan *outlier* dari *dataset*, 370 titik data akan ditinggalkan untuk analisis selanjutnya.

d. Transformasi Data

Tahap selanjutnya yaitu transformasi data dengan menggunakan *min-max normalization*. Data yang digunakan masih memiliki rentang yang berbeda-beda, maka untuk menghindari masalah numerik seperti *overflow* atau *underflow* perlu dilakukan normalisasi data.

Tabel 2 Data Sebelum Normalisasi

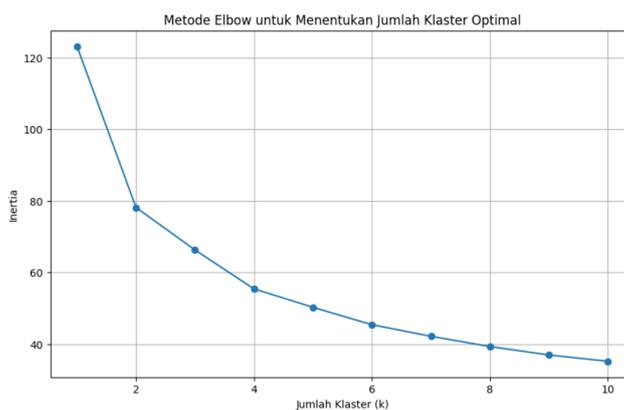
No.	Jumlah Penduduk	Usia Produktif (15-64)	...	Perdagangan	Pendidikan (SLTA ke Atas)	Perpindahan Penduduk
0	2280	1555	...	2	421	16
1	3381	2281	...	5	406	17
2	4006	2617	...	4	649	16
3	3124	2152	...	1	1195	21
4	3899	2646	...	3	1411	13

Tabel 3 Data Hasil Normalisasi

No.	Jumlah Penduduk	Usia Produktif (15-64)	...	Perdagangan	Pendidikan (SLTA ke Atas)	Perpindahan Penduduk
0	0.32	0.32	...	0.04	0.28	0.73
1	0.51	0.50	...	0.17	0.27	0.77
2	0.62	0.58	...	0.13	0.45	0.73
3	0.47	0.47	...	0.00	0.84	0.95
4	0.60	0.59	...	0.08	0.99	0.59

3.2.3 Clustering Menggunakan Metode K-Means

Proses clustering dalam penelitian ini dilakukan dengan metode K-Means dan menggunakan pendekatan Elbow untuk menentukan jumlah kluster yang paling optimal. Tahapan awal dilakukan dengan menerapkan algoritma K-Means untuk berbagai nilai k, mulai dari k = 1 sampai k = 10. Setiap kali clustering dijalankan, nilai *inertia* atau total jarak kuadrat dari data ke pusat kluster dihitung dan disimpan. Nilai-nilai *inertia* kemudian diplotkan dalam sebuah grafik sebagaimana yang terlihat pada gambar 2.



Gambar 2 Grafik Metode Elbow

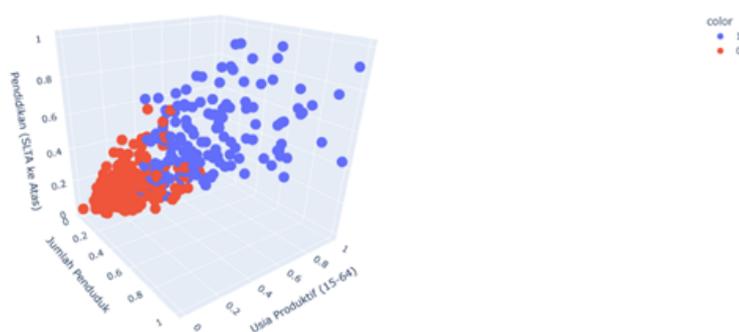
Berdasarkan Gambar 2, ada penurunan *inertia* yang tajam dari k = 1 ke k = 2, jumlah kluster optimal ditentukan pada k = 2, yang dikenal sebagai titik siku (*elbow*). Dengan k = 2, pengelompokan tetap sederhana dan efektif tanpa *overfitting*, memastikan data dikelompokkan dengan baik tanpa kompleksitas yang tidak diperlukan. Berikut ini tabel nilai *inertia* untuk setiap kluster:

Tabel 4 Inertia dan Penurunan Inertia

Jumlah Kluster (k)	Inertia	Penurunan Inertia	Persentase Penurunan (%)
1	123.18	NaN	NaN

Jumlah Klaster (k)	Inertia	Penurunan Inertia	Persentase Penurunan (%)
2	78.14	45.04	36.56
3	66.27	11.87	15.19
4	55.40	10.88	16.41
5	50.25	5.15	9.29
6	45.38	4.87	9.69
7	42.15	3.23	7.12
8	39.27	2.87	6.82
9	36.93	2.35	5.98
10	35.16	1.77	4.80

Setelah mengetahui jumlah klaster yang optimal, bahasa pemrograman python digunakan untuk membentuk model K-Means. Berikut ini visualisasi hasil pengelompokan data.



Gambar 3 Hasil Clustering

Gambar 3 menunjukkan hasil pengelompokan dengan teknik K-Means dengan dua klaster pada 370 titik data yang telah dibersihkan dari *outlier*. Tiga variabel yang ditampilkan untuk pengelompokan adalah jumlah penduduk, usia produktif (15-64 tahun), dan pendidikan (SMA atau lebih tinggi). Hasil pengelompokan menunjukkan pembagian data ke dalam dua kelompok yang ditandai dengan warna yang berbeda, di mana klaster 0 (merah) memiliki 223 data yang cenderung terkonsentrasi pada nilai yang lebih rendah untuk ketiga variabel yang ditampilkan, sedangkan klaster 1 (biru) terdiri dari 147 data dengan sebaran nilai yang lebih tinggi dan lebih tersebar pada ketiga dimensi variabel tersebut.

Tabel 5 Dataset Akhir

No	Kabupaten/ Kota	Kecamatan	Kelurahan/Desa	...	Potensi
1	Kab. Jeneponto	Tarowang	Tarowang	...	Tinggi
2	Kab. Jeneponto	Tarowang	Balangloe Tarowang	...	Tinggi
3	Kab. Jeneponto	Tarowang	Pao	...	Tinggi
4	Kab. Jeneponto	Tarowang	Bonto Ujung	...	Tinggi
5	Kab. Jeneponto	Batang	Bungeng	...	Tinggi

No	Kabupaten/ Kota	Kecamatan	Kelurahan/Desa	...	Potensi
6	Kab. Jeneponto	Arungkeke	Boronglamu	...	Rendah
...
368	Kab. Luwu Utara	Bone Bone	Tamuku	...	Rendah
369	Kab. Luwu Utara	Tana Lili	Sidomakmur	...	Rendah
370	Kab. Luwu Utara	Tana Lili	Munte	...	Tinggi

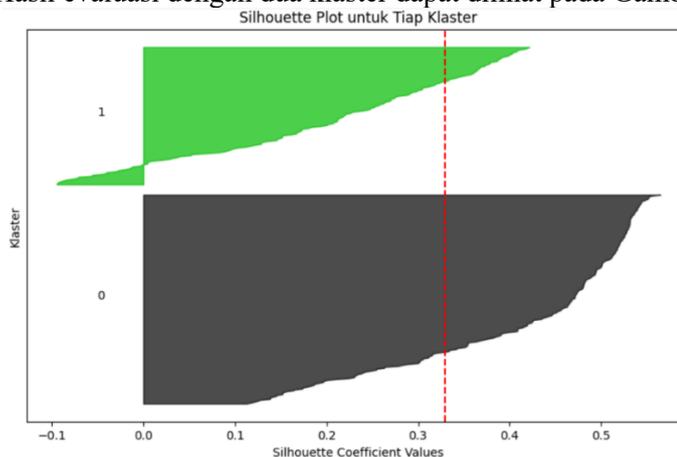
3.2.4 Hasil Evaluasi

Silhouette score digunakan untuk mengevaluasi kualitas kluster yang telah dibuat. Silhouette digunakan untuk melihat kualitas dan kekuatan kluster, mengukur seberapa baik atau buruknya suatu objek ditempatkan dalam suatu kluster. Nilai silhouette score berkisar antara -1 hingga 1, di mana nilai mendekati 1 menunjukkan bahwa objek dalam kluster sangat mirip satu sama lain dan sangat berbeda dari objek di kluster lain, yang berarti kluster yang dihasilkan memiliki struktur yang baik. Nilai mendekati 0 menunjukkan bahwa objek berada di atau sangat dekat dengan batas antara dua kluster, yang berarti kluster yang dihasilkan kurang terdefinisi dengan baik.

Tabel 6 Hasil Evaluasi Kluster

Kluster	Silhouette
2	0.33
3	0.26
4	0.26
5	0.25
6	0.24

Berdasarkan Tabel 6, semakin banyak kluster tidak meningkatkan kualitas pemisahan kelompok (seperti yang ditunjukkan oleh skor Silhouette yang menurun), mak pilihan terbaik adalah dua kluster karena hasil menunjukkan bahwa untuk $k = 2$, nilai Silhouette Score adalah 0.33, yang merupakan nilai tertinggi. Ini berbeda dengan jumlah kluster lain ($k=3$ hingga $k=6$) yang memiliki skor lebih rendah, yang berkisar antara 0.26 dan 0.24. Ini menunjukkan bahwa data yang terpisah dan batas antar-kluster lebih jelas dengan hanya dua kluster daripada dengan lebih banyak kluster. Hasil evaluasi dengan dua kluster dapat dilihat pada Gambar 4.



Gambar 4 Hasil Evaluasi Kluster

Pada Gambar 4, warna hijau menunjukkan klaster 1, dan warna hitam menunjukkan klaster 0. Garis merah putus-putus untuk hasil evaluasi, yang menunjukkan bahwa pengelompokan yang dilakukan memiliki skor *Silhouette* rata-rata sebesar 0,329, atau sekitar 0,33. Nilai ini menunjukkan bahwa kualitas pengelompokan berada di kisaran bawah hingga menengah, karena skor *Silhouette* bergerak dari -1 hingga 1, dengan nilai yang mendekati 1 mengindikasikan pengelompokan yang baik. Selain itu, terdapat 22 titik data dengan koefisien *Silhouette negatif*, yang mengindikasikan bahwa beberapa objek mungkin lebih cocok untuk ditempatkan di klaster lain karena kedekatannya dengan klaster terdekat.

3.2.5 Interpretasi Hasil

Berdasarkan hasil pengelompokan yang telah dilakukan ada dua kategori klaster dengan klaster tinggi dan rendah. Klaster pertama mencakup 60,27% kelurahan/desa, sedangkan klaster kedua mencakup 39,73% kelurahan/desa. Persentase setiap kriteria yang digunakan dalam analisis untuk menentukan klaster yang berpotensi tinggi atau rendah dapat dilihat pada Tabel 7.

Tabel 7 Persentase Klaster

Kriteria	C0 (%)	C1 (%)
Jumlah penduduk	36,99	63,01
Perpindahan penduduk	40,01	59,99
Usia produktif (15-64)	36,98	63,02
Tingkat pendidikan	35,79	64,21
Nelayan	55,63	44,37
Wiraswasta	32,23	67,77
Perdagangan	62,33	37,67

Berdasarkan Tabel 7, dapat disimpulkan bahwa klaster 1 (C1) merupakan klaster dengan potensi tinggi karena mendominasi indikator-indikator seperti jumlah penduduk, usia produktif, tingkat pendidikan, jumlah wiraswasta, dan mobilitas penduduk. Sementara itu, Klaster 0 (C0) dapat dikategorikan sebagai klaster berpotensi rendah meskipun unggul dalam sektor perdagangan dan kelautan, namun tertinggal dalam indikator-indikator pengembangan dinamika ekonomi secara keseluruhan.

4. KESIMPULAN

Berdasarkan penelitian yang dilakukan, wilayah pesisir Sulawesi Selatan dapat dikelompokkan menggunakan metode K-Means dengan mempertimbangkan variabel-variabel demografi seperti jumlah penduduk, tingkat pendidikan, usia produktif, dan sektor pekerjaan (jumlah nelayan, wiraswasta, dan perdagangan). Berdasarkan hasil yang diperoleh dari metode elbow, wilayah tersebut terbagi menjadi 2 klaster:

- Klaster pertama terdiri dari 223 kelurahan/desa dengan karakteristik yang memerlukan peningkatan keterampilan dalam pengelolaan sumber daya secara berkelanjutan.
- Klaster kedua mencakup 147 kelurahan/desa dengan karakteristik seperti jumlah nelayan, wiraswasta, dan pelaku perdagangan yang tinggi, serta tenaga kerja usia produktif dengan tingkat pendidikan yang relatif baik.

Evaluasi hasil pengelompokan daerah menggunakan silhouette score menghasilkan nilai sebesar 0,33. Berdasarkan hasil evaluasi tersebut, maka hasil pengelompokan yang terbentuk masuk dalam kategori struktur lemah. Selain itu, koefisien silhouette negatif (tumpang tindih) dua puluh dua titik data menunjukkan bahwa objek tertentu mungkin lebih baik ditempatkan di klaster lain karena mereka dekat dengan klaster terdekat.

5. SARAN

Adapun saran untuk penelitian selanjutnya:

- a. Ada beberapa data tumpang tindih, menurut hasil analisis. Sangat disarankan untuk mencoba teknik clustering alternatif atau menambahkan variabel tambahan untuk memperoleh pembagian data yang lebih baik.
- b. Diharapkan dapat memperluas cakupan wilayah dan menggunakan algoritma clustering dan tools yang berbeda untuk pengelompokan karakteristik penduduk.
- c. Penelitian ini hanya menggunakan rentang data Januari sampai Juni tahun 2024, diharapkan menggunakan dataset dengan rentang data lebih baru.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada BPS dan Direktorat Jenderal Kependudukan dan Pencatatan Sipil – Kementerian Dalam Negeri atas penyediaan data yang telah dipublikasi untuk digunakan sebagai sumber utama dalam analisis penelitian ini. Penulis juga mengucapkan terima kasih kepada Ibu Dr. Ir. Rismayani, S.Kom., M.T., selaku pembimbing I, serta Ibu Sriwahyuningsih Piu, S.Si., M.T., selaku pembimbing II, atas waktu dan pengetahuan yang telah diberikan dalam proses penyelesaian penelitian ini.

DAFTAR PUSTAKA

- [1] A. M. Cisneros-Montemayor dkk., “Enabling conditions for an equitable and sustainable blue economy,” *Nature*, vol. 591, no. 7850, hlm. 396–401, Mar 2021, doi: 10.1038/s41586-021-03327-3.
- [2] D. Aryanti, M. Zulkifli, N. Limayani, dan L. Retnosari, *Statistik Sumber Daya Laut dan Pesisir*, vol. 20. Badan Pusat Statistika, 2023.
- [3] Bidang Statistik Sosial, *Statistik Potensi Desa Provinsi Sulawesi Selatan*. Badan Pusat Statistik Provinsi Sulawesi Selatan, 2018.
- [4] K. H. Lee, J. Noh, dan J. S. Khim, “The Blue Economy and the United Nations’ sustainable development goals: Challenges and opportunities,” 1 April 2020, Elsevier Ltd. doi: 10.1016/j.envint.2020.105528.
- [5] Z. Xu, “Four major tasks of data science,” *Data Science and Management*, vol. 3, hlm. 1–2, Sep 2021, doi: 10.1016/J.DSM.2021.06.003.
- [6] H. Setiawan, “Pengembangan Aplikasi Sistem Bimbingan Tugas Akhir Berbasis Web Menggunakan Metode Rapid Application Development Pada Akademi Farmasi Cendikia Farma Husada,” 2024. Diakses: 16 Oktober 2024. [Daring]. Tersedia pada: <http://repo.darmajaya.ac.id/id/eprint/15152>
- [7] G. Triyandana, L. A. Putri, dan Y. Umaidah, “Penerapan Data Mining Pengelompokan Menu Makanan dan Minuman Berdasarkan Tingkat Penjualan Menggunakan Metode K-Means,” 2022. [Daring]. Tersedia pada: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [8] M. Maesaroh, T. Nur Padilah, dan J. Haerul Jaman, “Penerapan Algoritma K-Means Clustering Pada Pengelompokan Daerah Penyebaran Diare Di Provinsi Jawa Barat,” *Jurnal Mahasiswa Teknik Informatika*, vol. 7, no. 4, Jan 2023, doi: <https://doi.org/10.36040/jati.v7i4.7208>.
- [9] M. Khandava Mulyadien dan U. Enri, “Algoritma K-Means Untuk Pengelompokan Bantuan Langsung Tunai (BLT),” *Jurnal Ilmiah Wahana Pendidikan*, vol. 10, no. 20, hlm. 198–210, Nov 2024, doi: 10.5281/zenodo.6944517.
- [10] D. Syaputri, P. H. Noprita, dan S. Romelah, “Implementasi Algoritma K-Means untuk

- Pengelompokan Distribusi Sosial Ekonomi Masyarakat Berdasarkan Demografi Kependudukan,” vol. 1, hlm. 1–6, 2021.
- [11] Azrahwati, M. Nusrang, M. K. Aidid, dan Z. Rais, “K-Means Cluster Analysis for Grouping Districts in South Sulawesi Province Based on Village Potential,” *ARRUS Journal of Mathematics and Applied Science*, vol. 2, no. 1, hlm. 49–59, Feb 2022, doi: 10.35877/mathscience632.
- [12] Darmin, “Pendekatan CEPA (Collaborative, Ecology, Planning, Administrative) dalam Pemberdayaan Masyarakat Pesisir Pantai,” *Jurnal Sipatokkong BPSDM Sulawesi Selatan*, vol. 4, no. 2, Jun 2023.
- [13] A. A.-W. Karakara, J. A. Peparah, dan I. Dasmani, “Social resilience and demographic characteristics of coastal communities in Ghana: Implications for the blue economy,” 2023, doi: <http://dx.doi.org/10.2139/ssrn.4415011>.
- [14] H. Mukaromah, “Implementasi Blue Economy pada Sektor Kelautan di Surabaya,” UIN Sunan Ampel Surabaya, 2023. Diakses: 20 Februari 2025. [Daring]. Tersedia pada: <http://digilib.uinsa.ac.id/id/eprint/65328>
- [15] A. Sahjaya, R. Yusliandi, dan L. D. L. Siregar, “Tinjauan Literatur: Peranan Blue Economy dalam Meningkatkan Pendapatan Masyarakat Pesisir Melalui Sumber Daya Perikanan di Sumatera Utara,” *Jurnal Manuhara: Pusat Penelitian Ilmu Manajemen dan Bisnis*, vol. 3, no. 2, hlm. 17–27, Jan 2025, doi: <https://doi.org/10.61132/manuhara.v3i1.1590>.
- [16] A. Pangestu dan D. T. Ridwan, “Penerapan Data Mining Menggunakan Algoritma K-Means Pelanggan Berdasarkan Kubikasi Air terjual Menggunakan Weka,” 2021. [Daring]. Tersedia pada: <https://jurnal.umj.ac.id/index.php/just-it/index>