

Analisis Pengelompokan Tingkat Pengangguran dan Partisipasi Angkatan Kerja di Provinsi Indonesia

Abstrak

Pengangguran merupakan tantangan serius dalam konteks sosial dan ekonomi global, termasuk Indonesia. Dampaknya tidak hanya terbatas pada aspek ekonomi saja, namun juga mencakup aspek kesejahteraan yang dapat mempengaruhi Indeks Pembangunan Manusia (IPM). Penelitian ini bertujuan untuk menganalisis perubahan tingkat pengangguran dan pola partisipasi angkatan kerja di provinsi-provinsi Indonesia dengan menggunakan data Badan Pusat Statistik (BPS) pada bulan Februari hingga Agustus 2022. Konsep tingkat pengangguran terbuka (TPT) dan tingkat partisipasi angkatan kerja (TPAK) digunakan sebagai indikator penting untuk memahami tren pasar tenaga kerja. Pada penelitian ini menerapkan teknik K-means clustering untuk menganalisis data dan menggunakan metode elbow untuk menentukan jumlah cluster yang optimal. Hasil analisis memberikan gambaran rinci mengenai pola dinamis TPT dan TPAK di berbagai provinsi. Penelitian ini penting karena memberikan landasan kebijakan untuk mengurangi pengangguran dan meningkatkan tingkat partisipasi angkatan kerja di Indonesia. Hasil evaluasi menggunakan *silhouette coefficient* menunjukkan bahwa kualitas clustering baik, dan nilai $k=3$ mewakili jumlah cluster yang optimal. Analisis ini memberikan pemahaman komprehensif mengenai perbedaan dan kesenjangan antar provinsi dan memberikan wawasan berharga untuk pengembangan kebijakan ketenagakerjaan.

Kata kunci: Pengangguran, Clustering, K-Means

I. Pendahuluan

Pengangguran menjadi salah satu permasalahan sosial dan ekonomi yang senantiasa menjadi sorotan diberbagai negara, termasuk di Indonesia. Menurut Azzahra (2022), Pengangguran atau tunakarya merupakan orang yang tidak memiliki pekerjaan atau sedang berusaha melamar kerja. Menurut Badan Pusat Statistika, orang-orang yang sedang berjuang mencari profesi, sedang mempersiapkan usaha, ataupun orang-orang yang sudah memiliki pekerjaan, namun karena suatu hal sehingga masih mencari pekerjaan lainnya maka termasuk dalam kategori pengangguran terbuka.

Pengangguran bukan hanya sekedar fenomena ekonomi, namun juga memiliki dampak yang signifikan terhadap hal lainnya. Keterlibatan yang rendah dalam kegiatan ekonomi dapat merugikan, tidak hanya pendapatan individu tetapi juga memengaruhi kualitas kesejahteraan IPM. Menurut Muslim, M. R. (2014), terjadinya pengangguran di suatu negara dapat dikarenakan jumlah lapangan pekerjaan di suatu wilayah tertentu tidak dapat mencukupi jumlah Angkatan kerja atau adanya ketidakseimbangan antar jumlah permintaan lapangan kerja dengan penawaran yang tersedia. Pengangguran juga bisa diakibatkan oleh kualitas SDM (Sumber Daya Manusia) yang rendah dan tidak mampu untuk bersaing yang kemudian tersisih pada kompetisi pasar modern saat ini (Latifah dkk, 2017).

Menurut Munawir, dan Saharuddin. (2023). Tingkat Partisipasi Angkatan Kerja (TPAK) merupakan ukuran penyerapan tenaga kerja yang memberikan gambaran umum tentang keterlibatan penduduk dalam kegiatan ekonomi. Tingkat Partisipasi Angkatan Kerja (TPAK) menjadi cerminan dalam memahami interaksi antara sumber daya manusia (SDM) dan pasar tenaga kerja di suatu wilayah. Setiap provinsi di Indonesia memiliki dinamika yang unik dalam partisipasi Angkatan kerja, yang tercermin dalam kegiatan ekonomi dan sosial setempat. Peningkatan TPAK menjadi indikator positif yang menandakan keterlibatan yang lebih luas dari Angkatan kerja dalam aktivitas ekonomi. Begitupun sebaliknya, perubahan yang signifikan dalam Tingkat pengangguran dapat memengaruhi struktur TPAK secara menyeluruh.

Kondisi ini memberikan landasan untuk memahami bahwa pengangguran bukan hanya tantangan bagi pertumbuhan ekonomi biasa melainkan juga factor yang mempengaruhi tingkat kesejahteraan dan Pembangunan manusia secara menyeluruh. Dengan kondisi tersebut analisis tentang perubahan tingkat pengangguran dan pengelompokan tingkat partisipasi Angkatan kerja di berbagai provinsi Indonesia menjadi fokus utama untuk memahami dampak perubahan ekonomi dan sosial. Penelitian ini menggunakan pendekatan K-Means untuk menganalisis data Tingkat pengangguran dan partisipasi Angkatan kerja dari bulan februari sampai Agustus 2022.

Analisis ini menggunakan metode pengelompokan yang juga disebut clustering. Yang mana fungsi clustering menurut Bahtiar, D. (2023) adalah mengidentifikasi kelompok-kelompok dari suatu populasi dengan berbagai karakteristik khusus, clustering berbeda dengan classification, dimana pada clustering tidak ada definisi-definisi karakteristik awal yang telah ditentukan pada waktu classification. Bentuk clustering cukup bermacam, namun dalam analisis ini metode clustering yang akan di gunakan adalah metode K-Means. Metode K-Means clustering ini merupakan bentuk clustering yang umum digunakan dalam analisis data. Dalam konteks analisis pengelompokan tingkat pengangguran dan tingkat partisipasi Angkatan kerja, pendekatan K-Means diharap dapat memberikan pemahaman yang lebih mendalam tentang pola-pola dinamika TPT dan TPAK di setiap provinsi.

II. Kajian Pustaka

Pengangguran Terbuka

Dalam tulisan Franita dan Fuady, Menurut Sadono Sukirno (1994), menjelaskan bahwa pengangguran merupakan suatu keadaan di mana seseorang yang tergolong dalam angkatan kerja, ingin mendapatkan pekerjaan tetapi belum memperolehnya. Menurut Badan Pusat Statistik, menjelaskan bahwa pengangguran terbuka memiliki beberapa macam, yang pertama yaitu orang-orang yang tidak punya pekerjaan dan mencari pekerjaan, yang kedua adalah orang-orang yang tidak punya pekerjaan dan mempersiapkan usaha, kemudian yang ketiga merupakan orang-orang yang tidak punya pekerjaan dan tidak mencari pekerjaan, karena merasa tidak mungkin mendapat kerja, dan yang terakhir ialah orang-orang yang sudah punya pekerjaan tetapi belum mulai bekerja.

Penjelasan untuk konteks mempersiapkan suatu usaha atau pekerjaan tidak terbatas dalam jangka waktu seminggu yang lalu saja, tetapi bisa dilakukan beberapa waktu yang lalu asalkan seminggu yang lalu masih berusaha untuk mempersiapkan suatu kegiatan usaha.

Clustering

Dalam tulisan Edy Irwansyah, Menurut Tan (2006) menjelaskan bahwa clustering adalah sebuah proses untuk mengelompokan data ke dalam beberapa cluster atau kelompok yang ditentukan sehingga data dalam satu cluster memiliki Tingkat kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum. Sesuai dengan konsep dasar clustering bahwa clustering yang baik akan menghasilkan Tingkat kesamaan yang tinggi dalam satu kelas dan Tingkat kesamaan yang rendah untuk antar kelas. Bentuk nilai kesamaan antar kedua objek akan semakin tinggi ketika kedua objek tersebut memiliki kemiripan yang tinggi. Metode clustering secara umum dapat dibagi menjadi dua yaitu hierarki dan non-hierarki. Pada metode clustering non-hierarki jumlah kelompok ditentukan diawal sedangkan pada metode clustering hierarki, jumlah kelompok ditentukan di akhir analisis Anuraga, G. (2015).

Elbow Method

Metode Elbow adalah suatu metode yang digunakan untuk menentukan jumlah k optimal dari pembentukan sebuah cluster (Fuadah dkk, 2021). Metode Elbow merupakan salah satu metode dalam clustering untuk memberikan nilai atau jumlah cluster yang optimal berdasarkan perbandingan hasil persen antara jumlah cluster yang membentuk siku pada suatu titik. Metode elbow adalah pendekatan visual untuk menentukan jumlah kluster optimal pada algoritma K-Means. Pada metode ini, jumlah kluster diperoleh dengan mencari titik siku pada grafik yang menunjukkan penurunan jumlah kesalahan yang signifikan. Oleh karena itu, nilai k yang dipilih adalah nilai dimana penurunan kesalahan menjadi lebih lambat dan grafik membentuk sudut tajam, yang disebut elbow (Syakur dkk, 2018).

Perhitungan SSE (Sum of Square Error) adalah perhitungan yang biasa digunakan untuk mendapatkan perbandingan nilai k optimal dari masing-masing nilai cluster. Karena semakin besar jumlah k cluster maka nilai SSE akan semakin kecil. Berikut rumus SSE pada K-Means (Merliana dkk).

$$SSE = \sum_{K=1}^K \sum_{x_i \in S_K} \|X_i - C_k\|_2^2$$

dimana:

k = banyak cluster yang terbentuk, C_i adalah cluster ke-i, dan x merupakan data yang ada di setiap cluster. Dalam penjelasan Marliana dkk, dijelaskan bahwa setelah melakukan perhitungan SSE dengan k yang sudah inisialisasi dan menentukan titik mana penurunan SSE yang signifikan atau titik dimana grafik berbentuk siku maka pada nilai tersebutlah nilai k optimal berada dari perhitungan K-Means yang dilakukan.

K-Means Clustering

Algoritma K-Means adalah salah satu algoritma clustering non-hierarki paling populer. K-Means clustering adalah metode untuk menganalisa data atau data mining yang merupakan bentuk pemodelan unsupervised learning. Algoritma K-Means merupakan algoritma unsupervised yang mengambil sejumlah titik data memetakannya menjadi k cluster, jumlah cluster sesuai dengan k yang ditentukan yang mana dengan sejumlah k itulah juga centroid akan terbentuk (Shamrat dkk, 2020). K-Means merupakan algoritma yang membagi data menjadi beberapa cluster sehingga data serupa berada dalam cluster yang sama dan data berbeda berada dalam cluster berbeda. K-Means mempartisi data kedalam kelompok-kelompok yang mana semua data yang berada dalam satu kelompok memiliki karakteristik yang sama sedangkan untuk data antar kelompok memiliki nilai karakteristik yang berbeda. Algoritma K-Means ini memiliki tujuan untuk meminimalisasikan fungsi objektif yang telah di set dalam proses clustering, tujuan tersebut umumnya berusaha untuk meminimalkan variasi dalam satu kelompok dan memaksimalkan variasi antar kelompok (Gustientiedin, 2019).

Dalam tulisan Chandra dkk (2021), dijelaskan bahwa algoritma K-Means clustering memiliki nilai komputasi yang relatif cepat. Hal ini didasarkan terhadap analisis yang dilakukan oleh (Suomi G. dan Sanjay Kumar D), tentang perbandingan antara K-Means dan Fuzzy C-Means (FCM) yang mana hasil dari analisis tersebut membuktikan bahwa algoritma yang memiliki nilai kecepatan yang lebih unggul adalah algoritma K-Means dengan elapsed time 0.433755 detik dibandingkan dengan algoritma FCM yang nilai elapsedtime nya lebih besar dengan 0.781679 detik.

Silhouette Coefficient

Silhouette Coefficient merupakan bentuk gabungan dari metode cohesion dan separation. Yang mana tujuan dari cohesion yaitu mengukur seberapa dekat hubungan antara objek salah satu kelompok cluster, dan tujuan dari separation ialah mengukur jarak antar cluster dengan cluster lainnya. Jadi *silhouette*

coefficient merupakan suatu cara yang digunakan untuk mengukur seberapa baik sebuah objek diklasifikasikan saat menggunakan algoritma clustering. Nilai cluster terbaik menurut perhitungan Silhouette coefficient ini adalah nilai dengan rata-rata mendekati satu (Simanjuntak dan Khaira, 2021). Berikut rumus Silhouette Coefficient dalam tulisan Rousseeuw (1987):

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

dimana: $a(i)$ adalah jarak antara objek data i dengan cluster tempat objek i berada yang berbeda, sedangkan $b(i)$ adalah jarak antara objek data i dengan cluster tetangga. Nilai $s(i)$ juga bisa didapat dengan beberapa ketentuan seperti berikut:

$$s(i) = \begin{cases} 1 - \frac{a(i)}{b(i)}, & \text{if } a(i) < b(i) \\ 0, & \text{if } a(i) = b(i) \\ b(i)/a(i) - 1, & \text{if } a(i) > b(i) \end{cases}$$

Ketika cluster A hanya berisi satu objek dan membuat definisi nilai $a(i)$ tidak jelas. Oleh karena itu maka ditetapkan untuk $s(i)$ sama dengan nol. Pilihan ini dapat sewenang-wenang namun nilai nol tampaknya nilai paling netral dalam konteks ini. Dengan kata lain, ketika suatu cluster hanya terdiri dari satu objek, maka susah untuk menentukan seberapa baik atau buruknya objek tersebut. Oleh karena dipilihlah nilai yang dianggap paling netral untuk nilai $s(i)$ yaitu nol.

III. Metode

Data Penelitian

Penelitian dilakukan dengan metode literatur review. Pada tahap ini dilakukan studi literatur dengan mencari bahan referensi berupa buku, jurnal, dan Internet sesuai dengan permasalahan yang dibahas. Data yang digunakan dalam artikel ini merupakan data sekunder Tingkat pengangguran yang dikelola oleh Badan Pusat Statistik Indonesia (BPS). Data tersebut dapat diakses melalui website resmi BPS dengan kata kunci pencarian Tingkat pengangguran. Pada data tersebut terdiri dari data Tingkat Pengangguran Terbuka (TPT) dan Tingkat Partisipasi Angkatan Kerja (TPAK) Menurut Provinsi, pada tahun 2022.

Praproses data

Praproses data merupakan tahap awal dalam analisis data yang bertujuan untuk memastikan bahwa data yang akan dianalisis merupakan data yang sudah siap untuk diolah. Praproses data ini mencakup memahami tipe data, kemudian handling missing values untuk mengatasi nilai-nilai yang hilang dalam data set, pembersihan data untuk membuang data yang mungkin tidak digunakan, handling categorical data untuk mengubah variable kategori menjadi bentuk numerik sehingga dapat diolah oleh algoritma yang akan digunakan, dan analisis statistik untuk mengetahui karakteristik data. Tujuan utama praproses data ini adalah untuk meningkatkan kualitas data dan mempermudah untuk interpretasi.

Algoritma K-Means

K-Means adalah metode clustering berbasis jarak yang membagi data kedalam sejumlah cluster dan algoritma ini hanya bekerja pada atribut numerik. Algoritma K-Means merupakan metode non-hierarki

yang awalnya mengambil sebagian banyaknya komponen populasi untuk dijadikan pusat cluster awal (Metisen dan Sari, 2015).

Langkah-langkah melakukan clustering dengan metode K-Means menurut Sulistiyawati dan Supriyanto (2023), yaitu:

1. Menentukan nilai k sebagai jumlah cluster yang ingin dibentuk.
2. Inisialisasi k pusat cluster atau centroid Langkah ini ketika diawal biasanya dilakukan secara random
3. Menghitung jarak setiap data input terhadap masing-masing centroid menggunakan rumus jarak Euclidean (Euclidean Distance) hingga menemukan jarak paling dekat dari setiap data dengan centroid. Berikut persamaan Euclidean Distance:

$$De = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

dimana:

De adalah *Euclidean Distance*, i merupakan banyaknya objek, (x,y) merupakan koordinat objek dan, (s,t) merupakan koordinat centroid.

4. Mengklasifikasikan setiap data berdasarkan kedekatannya dengan centroid (jarak terkecil).
5. Memperbarui perbaharui nilai centroid, yang mana nilai centroid yang baru didapat dari rata-rata cluster yang bersangkutan menggunakan rumus berikut:

$$v_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

dimana:

v_{ij} merupakan centroid atau rata rata cluster ke-i untuk variable ke-j, lalu N_i merupakan jumlah data yang menjadi anggota cluster ke-i, (i,k) adalah indeks cluster, j adalah indeks variable, dan X_{kj} adalah nilai data ke-k yang ada di dalam cluster tersebut untuk variable ke-j.

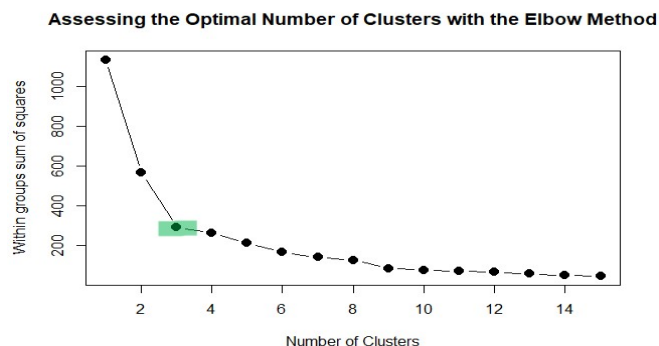
6. Melakukan perulangan dari Langkah 2 hingga 5, sampai anggota setiap cluster tidak ada yang berubah.

IV. Hasil dan Pembahasan

Dalam mengolah data perlu diketahui terlebih dahulu karakteristik data mentahnya sebelum diolah. Data yang telah dikumpulkan adalah data tingkat pengangguran dan tingkat partisipasi Angkatan kerja bulan Februari dan Agustus 2022. Data berjumlah 35 baris dengan 5 kolom, data awal ini dapat dilihat pada tabel 1. Subjek clustering ini adalah provinsi-provinsi di Indonesia, yang mana atributnya adalah tingkat pengangguran dan tingkat partisipasi Angkatan kerja. Sebelum menganalisis data menggunakan metode yang lebih spesifik, sangat perlu untuk memahami karakteristik awal data menggunakan analisis statistik deskriptif, untuk mendapatkan gambaran ringkas dan informatif tentang karakteristik dasar suatu dataset. Karakteristik dari analisis statistik didapatkan informasi tentang sebaran nilai, kecenderungan pusat, dan variasi dari masing-masing variable dalam data. Berdasarkan hasil analisis statistik didapatkan bahwa Tingkat Pengangguran Terbuka (TPT) memiliki rata-rata tertinggi pada bulan Februari dengan nilai 5.14 dan terendah pada bulan Agustus dengan nilai 4.99. Untuk Tingkat Partisipasi Angkatan Kerja (TPAK) memiliki rata-rata tertinggi pada bulan Februari dengan nilai 68.78 dan terendah pada bulan Agustus dengan nilai 68.64. Setelah memahami dan mempersiapkan data dengan memeriksa nilai yang kosong dan menganalisis statistik maka data sudah siap untuk lanjut kelangkah berikutnya.

Langkah pertama clustering K-Means adalah menentukan nilai k sebagai jumlah cluster yang akan dibentuk. Metode yang akan digunakan untuk menentukan nilai k optimal dalam artikel ini adalah Elbow

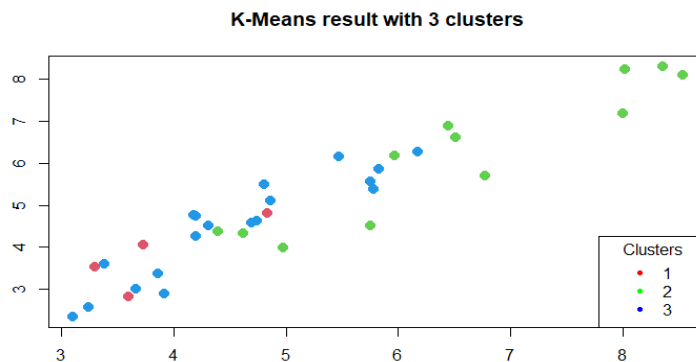
Method. Perhitungan Elbow dalam artikel ini menggunakan bantuan software R Studio yang nantinya juga dapat menghasilkan plot hasil untuk menunjukkan titik mana k cluster optimal berada. Rentang iterasi yang dilakukan untuk mendapatkan nilai optimal k adalah dari rentang 2 hingga 15. Hasil plot elbow pada Gambar 1 menunjukkan bahwa nilai optimal k berada di cluster 3 dengan tanda siku sebagai tandanya, maka nilai k optimalnya adalah 3.



Gambar 1. Hasil jumlah cluster optimal menggunakan *Elbow Method*.

Langkah selanjutnya adalah melakukan proses clustering menggunakan metode K-Means dengan jumlah cluster k yang sudah ditentukan dari hasil Elbow method pada data tingkat pengangguran dan tingkat partisipasi angkatan kerja yang akan di cluster. Algoritma K-Means diterapkan untuk mengelompokkan data ke dalam k cluster yang akan memberikan hasil berupa pengelompokan dengan karakteristik yang berbeda pada setiap cluster data TPT dan TPAK.

Saat melakukan pengelompokan ini, memperhitungkan jarak antara titik data dan pusat cluster, sehingga semua data ditugaskan ke cluster yang paling dekat dengan pusatnya (Sulistiyawati dan Supriyanto, 2023). Pada akhirnya, data tersebut dikelompokkan ke dalam cluster-cluster yang mencerminkan pola atau persamaan tertentu pada data TPT dan TPAK. Proses ini membantu kita memahami perbedaan dan kesenjangan dalam tingkat pengangguran dan partisipasi angkatan kerja di berbagai provinsi. Hasil pengelompokan menggunakan K-Means clustering dengan jumlah k cluster 3, terpresentasikan pada Gambar 2.



warna merah, data cluster 2 dilambangi dengan warna hijau, dan data di cluster 3 dilambangi dengan warna biru.

CLUSTER LABEL	TPT- FEBRUARI	TPT- AGUSTUS	TPAK- FEBRUARI	TPAK- AGUSTUS	BANYAK DATA
1	3.867500	3.807500	76.42750	75.61000	4
2	6.528333	6.197500	64.88500	65.18333	12
3	4.536842	4.479474	69.62632	69.36211	19

Tabel 2. Hasil rata-rata Clustering data TPT dan TPAK

Hasil rata-rata pada tabel 2, menunjukan karakteristik setiap clustering. Dalam hasil tersebut terlihat bahwa tingkat pengangguran pada bulan Agustus mengalami penurunan dari bulan Februari pada semua cluster. Tetapi untuk TPAK pada setiap cluster berbeda ada yang mengalami penurunan dan ada yang mengalami peningkatan walaupun tidak signifikan. Pada cluster 1 yang terdiri dari 4 provinsi yang terdiri atas provinsi DI Yogyakarta, Bali, Nusa Tenggara Timur, dan Papua memiliki tingkat pengangguran (TPT) yang paling rendah dari cluster lainnya dan untuk nilai TPAK pada cluster1 meru pakan nilai tertinggi daripada cluster lainnya walaupun mengalami penurunan pada bulan Agustus.

Pada cluster 2 terdapat 12 provinsi yang termasuk dalam cluster tersebut diantaranya terdapat provinsi Aceh, Riau, Kepulauan Riau, DKI Jakarta, Jawa Barat, Banten, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Selatan, Maluku, Maluku Utara. Pada cluster 2 ini memiliki karakteristik dengan tingkat pengangguran (TPT) yang lebih tinggi dibandingkan dengan cluter 1, TPAK pada cluster ini memiliki nilai yang lebih rendah dari cluster 1, namun TPAK pada cluster 2 ini mengalami peningkatan dari bulan Februari hingga bulan Agustus yang mana menunjukkan bahwa pada provinsi-provinsi yang tergolong dalam cluster 2 ini mengalami perkembangan yang baik.

Pada cluster 3 terdapat 19 provinsi yang terdiri dari provinsi Sumatera Utara, Sumatera Barat, Sumatera Selatan, Jambi, Bengkulu, Lampung, Kepulauan Bangka Belitung, Jawa Tengah, Jawa Timur, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Papua Barat, dan Indonesia pada cluster ini tingkat pengangguran (TPT) berada di tengah-tengah dari cluster lainnya, nilai TPAK pada cluster ini juga lebih berada di tengah-tengah dari nilai cluster lainnya dan nilai TPAK pada cluster ini mengalami sedikit penurunan pada bulan Agustusdari bulan Februari.

Dengan demikian, dapat diinterpretasikan bahwa cluster 1 merupakan provinsi dengan karakteristik tingkat TPT yang rendah dan tingkat TPAK yang tinggi. Lalu cluster 2 merupakan provinsi dengan tingkat TPT yang tinggi dan tingkat TPAK yang rendah. Cluster 3 merupakan provinsi-provinsi dengan tingkat TPT dan TPAK sedang. Setelah dilakukan proses clustering menggunakan K-Means, Langkah selanjutnya adalah melakukan *Silhouette coefficient* untuk mengukur seberapa baik objek di klasifasikan dalam cluster. Dengan ketentuan nilai diantara 1 sampai -1, yang mana nilai yang semakin mendekati 1 maka semakin baik kualitas objek tersebut dalam cluster, begitu pula sebaliknya. Hasil *Silhouette coefficient* dapat dilihat pada Tabel 3.

CLUSTER LABEL	BANYAK PROVINSI	RATA-RATA LEBAR SILHOUETTE
1	4	0.4212400
2	12	0.3933311
3	19	0.4840267

Tabel 3. Hasil *Silhouette Coefficient*

Tabel 3 menunjukkan hasil dari penggunaan silhouette coefficient pada analisis clustering ini. Nilai rata-rata silhouette yang positif pada semua cluster menunjukkan bahwa pengelompokan yang dihasilkan oleh algoritma K-Means memiliki kualitas yang baik. Tingginya nilai pada hasil silhouette coefficient ini menandakan bahwa objek-objek dalam masing-masing cluster cenderung serupa dari pada dengan objek di cluster lain. Hal ini konsisten dengan hasil K-Means sebelumnya bahwa setiap cluster memiliki tingkat atau karakteristik yang berbeda.

Dengan demikian, hasil analisis clustering dan evaluasi Silhouette coefficient ini diharapkan dapat menjadi tujuan baik sebagai landasan kuat untuk memahami dan merumuskan kebijakan di tingkat provinsi terkait tingkat pengangguran (TPT) dan tingkat angkatan partisipasi kerja (TPAK), yang mana diharapkan dapat menjadi dasar untuk memandu keputusan strategis dalam mengatasi tantangan tingkat pengangguran dan memperkuat tingkat partisipasi angkatan kerja di seluruh provinsi Indonesia.

V. Kesimpulan

Pada artikel ini, analisis dilakukan menggunakan metode clustering K-means untuk menganalisis data tingkat pengangguran dan partisipasi tenaga kerja di provinsi-provinsi di Indonesia. Sebelumnya, karakteristik data awal diidentifikasi melalui analisis statistik deskriptif, sehingga dapat diketahui bahwa bulan Februari memiliki rata-rata TPT tertinggi dan bulan Agustus memiliki TPAK tertinggi. Selanjutnya dilakukan proses clustering K-means dengan mencari nilai k optimal menggunakan metode elbow sehingga diperoleh nilai k adalah 3.

Hasil clustering menunjukkan tiga cluster dengan sifat berbeda. Cluster 1 yang terdiri dari provinsi seperti DI Yogyakarta, Bali, Nusa Tenggara Timur, dan Papua memiliki tingkat TPT terendah dan tingkat TPAK tertinggi, meski tercatat mengalami penurunan pada bulan Agustus. Cluster 2, yang mencakup provinsi seperti Aceh, Riau, dan Jakarta Metropolitan, memiliki tingkat TPT tinggi dan tingkat TPAK rendah, dengan peningkatan TPAK dari bulan Februari hingga Agustus. Cluster 3 yang mencakup provinsi seperti Jawa Tengah, Jawa Timur, dan Sulawesi Selatan memiliki nilai TPT dan TPAK sedang, meskipun terdapat sedikit penurunan pada bulan Agustus.

Berdasarkan evaluasi dengan silhouette coefficient menunjukkan kualitas pengelompokan yang baik dengan setiap cluster bernilai positif. Secara ringkas, analisis ini memberikan pemahaman komprehensif mengenai perbedaan dan kesenjangan tingkat pengangguran dan partisipasi angkatan kerja antar provinsi di Indonesia. Hasil penelitian ini dapat memberikan dasar untuk mengembangkan langkah-langkah strategis untuk mengatasi masalah pengangguran dan memperkuat partisipasi buruh di tingkat negara bagian.

Daftar Pustaka

- Muslim, M. R. (2014). Pengangguran Terbuka dan Determinannya. *Jurnal Ekonomi dan Studi Pembangunan*, 15(2), 171-181. Institute of Public Policy and Economic Studies (INSPECT) Yogyakarta, Jalan Kenari R 13 Sidoarum III, Godean, Sleman, Yogyakarta, 55564 Indonesia.
- Latifah, N., Rotinsulu, D. C. H., & Tumilaar, R. L. H. (2017). Pengaruh Pertumbuhan Ekonomi dan Indeks Pembangunan Manusia Terhadap Tingkat Pengangguran Terbuka dan Dampaknya pada Jumlah Penduduk Miskin di Kota Manado [Effect of Economic Growth and Human Development Index on Unemployment Rate and Number of Poor People in Manado City]. *Jurnal Berkala Ilmiah Efisiensi*, 17(02).
- Azzahra, D. D. G., Aini, W. R., & Desmawan, D. (2022). Analisis Dampak Tingkat Pengangguran Terbuka terhadap Kemiskinan Menurut Kabupaten dan Kota Selama Pandemi COVID-19 di Provinsi Banten. *Profit: Jurnal Manajemen, Bisnis dan Akuntansi*, 1(4), 01-09. <https://doi.org/10.58192/profit.v1i4.174>
- Munawir, & Saharuddin. (2023). Pengaruh Partisipasi Angkatan Kerja (TPAK) dan Tingkat Pengangguran Terbuka (TPT) terhadap PDRB di Provinsi Aceh. *Jurnal Ekonomika Indonesia*, 12(1). <https://ojs.unimal.ac.id/index.php/ekonomika>
- Metisen, B. M., & Sari, H. L. (2015). Analisis Clustering Menggunakan Metode K-Means dalam Pengelompokan Penjualan Produk pada Swalayan Fadhila. *Jurnal Media Infotama*, 11(2), 110. ISSN 1858-2680. Fakultas Ilmu Komputer, Universitas Dehasen Bengkulu.
- Sulistiyawati, A., & Supriyanto, E. (2023). Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan. *Jurnal TEKNO KOMPAK*, 15(2), 25-36. P-ISSN: 1412-9663, E-ISSN: 2656-3525. Fakultas Teknik dan Ilmu Komputer, Sistem Informasi, Universitas Teknokrat Indonesia, Bandar Lampung, Indonesia.
- Bahtiar, D. (2023). Pemetaan Penduduk Penerima Bantuan Sosial Desa Waru Jaya Menggunakan Algoritma K-Means Clustering. *Scientia Sacra: Jurnal Sains, Teknologi dan Masyarakat*, 3(2), 29.
- Irwansyah, E. Clustering. *BINA NUSANTARA School of Computer Science*. Retrieved from <https://socs.binus.ac.id/2017/03/09/clustering/>
- Anuraga, G. (2015). Hierarchical Clustering Multiscale Bootstrap untuk Pengelompokan Kemiskinan di Jawa Timur. *Jurnal Statistika*, 3(1). Retrieved from <https://jurnal.unimus.ac.id/index.php/statistik/article/view/1432/1485>
- Franita, R., & Fuady, A. (2019). Analisa Pengangguran di Indonesia. *Nusantara: Jurnal Ilmu Pengetahuan Sosial*, 2, 88. ISSN 2541-657X.
- Badan Pusat Statistik. Tenaga Kerja. Diakses dari <https://www.archive.bps.go.id/subject/6/tenaga-kerja.html#subjekViewTab1>
- Fuadah, A. W., Arifin, F. N., & Juwita, O. (2021). Optimasi K-Klasterisasi Ketahanan Pangan Kabupaten Jember Menggunakan Metode Elbow. *Informatics Journal*, 6(3), 136. ISSN: 2503-250X.
- Merliana, N. P. E., Ernawati, & Santoso, A. J. Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means Clustering. Dalam *Prosiding Seminar Nasional Multi Disiplin Ilmu*. ISBN: 978-979-3649-81-8.
- Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration K-Means Clustering Method and Elbow Method for Identification of The Best Customer Profile Cluster. *IOP Conf. Series: Materials Science and Engineering*, 336(1), 012017. doi:10.1088/1757-899X/336/1/012017.

- Shamrat, F. M. J. M., Tasnim, Z., Mahmud, I., Jahan, M. N., & Nobel, N. I. (2020). Application Of K-Means Clustering Algorithm To Determine The Density Of Demand Of Different Kinds Of Jobs. *International Journal of Scientific & Technology Research*, 9(02), 2550. ISSN 2277-8616. [www.ijstr.org]
- Chandra, M.D., Irawan, E., Saragih, I.S., Windarto, A.P. and Suhendro, D. (2021). Penerapan Algoritma K-Means dalam Mengelompokkan Balita yang Mengalami Gizi Buruk Menurut Provinsi. *BIOS: Jurnal Teknologi Informasi dan Rekayasa Komputer*. 2, 1 (Mar. 2021), 30-38. DOI:<https://doi.org/10.37148/bios.v2i1.19>.
- Gustientiedina, M., Hasnil Adiya, & Yenny Desnelita. (2019). Penerapan Algoritma K-Means untuk Clustering Data Obat-Obatan pada RSUD Pekanbaru. *Jurnal Nasional Teknologi dan Sistem Informasi*, 5(1), 17-24. Retrieved from [http://teknosi.fti.unand.ac.id/Jurnal Nasional Teknologi dan Sistem Informasi/vol.05no.01/2019/017-024](http://teknosi.fti.unand.ac.id/Jurnal_Nasional_Teknologi_dan_Sistem_Informasi/vol.05no.01/2019/017-024). doi:10.25077/TEKNOSI.v5i1.2019.17-24
- Simanjuntak, K. P., & Khaira, U. (2021). Hotspot Clustering in Jambi Province Using Agglomerative Hierarchical Clustering Algorithm. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), 7-16. Retrieved from <https://journal.irpi.or.id/index.php/malcom/article/view/6>. P-ISSN: 2797-2313, E-ISSN: 2775-85757.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53–65. doi:10.1016/0377-0427(87)90125-7.

Lampiran

Tabel 1. Data awal dari BPS

PROVINSI	TINGKAT PENGANGGURAN TERBUKA (TPT)- FEBRUARI	TINGKAT PENGANGGURAN TERBUKA (TPT)- AGUSTUS	TINGKAT PARTISIPASI ANGKATAN KERJA (TPAK)- FEBRUARI	TINGKAT PARTISIPASI ANGKATAN KERJA (TPAK)- AGUSTUS
ACEH	5.97	6.17	64.15	63.5
SUMATERA UTARA	5.47	6.16	70.45	69.53
SUMATERA BARAT	6.17	6.28	68	69.3
RIAU	4.4	4.37	65.98	63.86
JAMBI	4.7	4.59	68.86	67.84
SUMATERA SELATAN	4.74	4.63	69.33	69.31
BENGKULU	3.39	3.59	71.49	69.81
LAMPUNG	4.31	4.52	72.51	70.06
KEPULAUAN BANGKA BELITUNG	4.18	4.77	70.16	67.38
KEPULAUAN RIAU	8.02	8.23	66.48	68.94
DKI JAKARTA	8	7.18	62.27	63.08
JAWA BARAT	8.35	8.31	66.31	66.15
JAWA TENGAH	5.75	5.57	72.04	70.84
DI YOGYAKARTA	3.73	4.06	74.68	72.6
JAWA TIMUR	4.81	5.49	70.99	71.23
BANTEN	8.53	8.09	65.08	64.72
BALI	4.84	4.8	77.14	76.86
NUSA TENGGARA BARAT	3.92	2.89	70.25	70.93
NUSA TENGGARA TIMUR	3.3	3.54	73.66	75.23
KALIMANTAN BARAT	4.86	5.11	69.74	68.97
KALIMANTAN TENGAH	4.2	4.26	66.94	67.23
KALIMANTAN SELATAN	4.2	4.74	69.64	67.55
KALIMANTAN TIMUR	6.77	5.71	66.22	64.73
KALIMANTAN UTARA	4.62	4.33	64.52	67.62
SULAWESI UTARA	6.51	6.61	61.97	63.08
SULAWESI TENGAH	3.67	3	70.86	69.99
SULAWESI SELATAN	5.75	4.51	65.85	66.18
SULAWESI TENGGARA	3.86	3.36	66.31	68.82
GORONTALO	3.25	2.58	67.78	68.91
SULAWESI BARAT	3.11	2.34	68.9	73
MALUKU	6.44	6.88	63.08	65.46
MALUKU UTARA	4.98	3.98	66.71	64.88
PAPUA BARAT	5.78	5.37	69.59	68.55
PAPUA	3.6	2.83	80.23	77.75
INDONESIA	5.83	5.86	69.06	68.63