



## ***Clustering of Districts and Cities in North Sumatra Province Based on People's Welfare Indicators Using K-Means Algorithm***

### **Pengelompokan Kabupaten dan Kota Provinsi Sumatera Utara Berdasarkan Indikator Kesejahteraan Rakyat Menggunakan Algoritma K-Means**

**Yustina Lentiana Dakhi<sup>1</sup>, Besse Arnawisuda Ningsi<sup>2\*</sup>**

<sup>1,2</sup>Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam (MIPA)  
Universitas Pamulang, Indonesia

E-Mail: <sup>1</sup>yustinalentianadakhidakhi@gmail.com, <sup>2</sup>dosen00205@unpam.ac.id

*Received Apr 08th 2024; Revised Jun 15th 2024; Accepted Jun 24th 2024*  
*Corresponding Author: Besse Arnawisuda Ningsi*

#### **Abstract**

*The welfare of the people is the main goal of the Indonesian state which is enshrined in the Constitution (1945) and the fifth principle of Pancasila. North Sumatra is a province in Indonesia which is located in the northern part of the island of Sumatra. The island has 33 districts/cities, consisting of 25 prefectures and 8 cities. The aim of this research is to determine the results of the grouping of Regencies/Cities of North Sumatra Province based on People's Welfare Indicators using the k-means cluster method. The variables observed in this research are population, health, education, employment, tariffs and consumption patterns and poverty. The results of the research in grouping the characteristics of clusters formed based on indicators of people's welfare are Cluster 1 consisting of North Tapanuli, Toba Samosir, Dairi, Humbang Hasundutan, Pakpak Bharat, Samosir, West Nias. And Cluster 2 consists of Mandailing Natal, South Tapanuli, Central Tapanuli, Labuhan Batu, Asahan, Simalungun, Karo, Deli Serdang, Langkat, South Nias, Serdang Bedagai, Batu Bara, North Padang Lawas, Padang Lawas, South Labuhanbatu, North Labuhanbatu, North Nias, Tanjungbalai, Gunungsitoli Cluster 3 consists of Sibolga, Pematangsiantar, Tebing Tinggi, Medan, Binjai, Padangsidimpuan. Meanwhile, Cluster 4 consists of Nias. The conclusion of this research can be seen from the Silhouette Index test results showing that forming 4 clusters has the best accuracy value compared to forming 2 clusters and forming 3 clusters.*

*Keyword: Grouping, Indicator, K-means Clustering, People's Welfare*

#### **Abstrak**

Secara bahasa umrah bermakna ziarah atau berkunjung, sedangkan secara istilah umrah adalah perjalanan ke Baitullah di luar waktu haji dengan tujuan melaksanakan ibadah tertentu dan memenuhi syarat-syarat khusus. PT Hajar Aswad merupakan sebuah perusahaan travel umrah yang beroperasi di Indonesia. PT Hajar Aswad bertanggung jawab untuk mengatur perjalanan, akomodasi, transportasi, dan berbagai keperluan lainnya bagi para jemaah umrah, untuk itu perlu memiliki pemahaman yang baik mengenai pola dan tren jumlah jemaah umrah agar dapat mengoptimalkan operasional dan memberikan pelayanan yang memuaskan kepada jemaah. Oleh karena itu penelitian ini dilakukan untuk memprediksi jumlah jemaah umrah pada PT Hajar Aswad menggunakan algoritma RNN dan LSTM agar PT Hajar Aswad. Hasil perbandingan kedua algoritma menunjukkan bahwa LSTM mampu memberikan hasil prediksi yang sedikit lebih baik dibandingkan RNN dengan parameter window size 7, optimizer Adam, batch size 8, dan learning rate 0,01. Model ini memiliki nilai RMSE sebesar 0,1758, MAPE sebesar 0,4846, dan R2 sebesar 0,5198.

**Kata Kunci:** Indikator, Kesejahteraan Rakyat, K-Means Clustering, Pengelompokan

#### **1. PENDAHULUAN**

Kesejahteraan rakyat merupakan tujuan utama negara Indonesia yang diabadikan dalam Undang-undang Dasar (UUD) 1945 dan sila kelima Pancasila. Masyarakat adalah subjek dan objek, yang merupakan peran mendasar dari kesejahteraan. Manusia adalah kekayaan negara atau bangsa, oleh karena itu masyarakat adalah objek dan agen utama kesejahteraan. Kesejahteraan adalah keadaan yang baik, keadaan masyarakat yang sejahtera, sehat dan damai. Dalam kebijakan sosial, kesejahteraan mengarah pada pelayanan untuk

memenuhi kebutuhan masyarakat. Kesejahteraan komunitas sosial tergantung pada kepuasan kebutuhan primer, sekunder dan tersier.

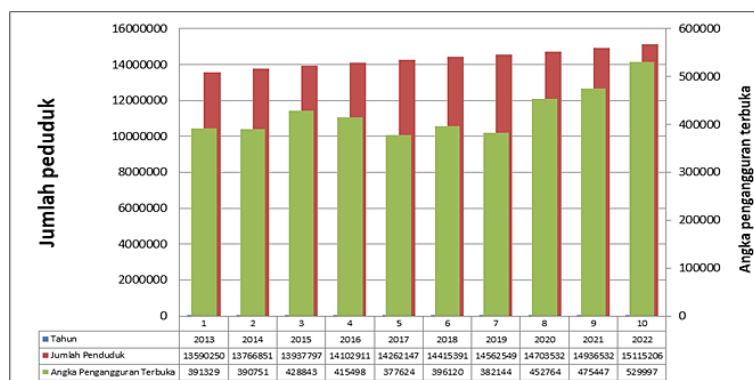
Kesejahteraan rakyat pada dasarnya merupakan suatu kondisi yang bentuknya dinamis atau dengan kata lain nilai kuantitatifnya tidak akan pernah berhenti karena akan terus berubah seiring dengan perkembangan kebutuhan hidup manusia [1]. Permasalahan kesejahteraan sosial yang berkembang saat ini menunjukkan adanya warga negara yang hak-hak dasarnya tidak terpenuhi dengan baik karena tidak mendapat dukungan pendapatan dari negara. Akibatnya, warga terus menghadapi kendala dalam mengikuti kegiatan sosial sehingga tidak bisa hidup layak.

Kesejahteraan masyarakat di bidang sosial pada dasarnya merupakan keadaan sosial yang memungkinkan bagi setiap warga Negara untuk dapat memenuhi kebutuhan hidup yang bersifat jasmani, rohani dan sosial sesuai dengan hakekat dan martabat manusia untuk dapat mengatasi berbagai masalah sosial yang dihadapi diri, keluarga dan masyarakatnya untuk berkembang menjadi lebih baik [2]. Konsep kesejahteraan dapat dibedakan menjadi kesejahteraan individu dan kesejahteraan sosial. Kesejahteraan individu merupakan cara mengaitkan kesejahteraan dengan pilihan individu secara obyektif, pilihan yang dilakukan individu sebagai uji yang obyektif adalah membandingkann kesejahteraan individu pada situasi yang berbeda kesejahteraan sosial merupakan cara mengaitkan kesejahteraan dengan cara menjumlahkan kepuasan seluruh individu dalam masyarakat [3].

Menurut Bintarto kesejahteraan dapat diukur dari aspek kehidupan Dengan melihat kualitas hidup dari segi materi seperti kualitas rumah, bahan pangan, dan sebagainya. Dengan melihat kualitas hidup dari segi fisik seperti kesehatan tubuh, lingkungan alam, dan sebagainya. Dan Dengan melihat kualitas hidup dari segi mental seperti fasilitas pendidikan, lingkungan budaya, dan sebagainya [4].

Sumatera Utara merupakan salah satu provinsi di Indonesia yang terletak di bagian utara pulau Sumatera. Di pulau tersebut memiliki 33 kabupaten/kota, terdiri dari 25 prefektur dan 8 kota. Ada beberapa suku bangsa seperti Batak Toba, Karo, Pakpal-Dairi, Simalungun, Mandailing, Melayu dan Nias. Dalam setiap sukunya memiliki budaya dan tradisi yang berbeda. Provinsi Sumatera Utara memiliki luas 72.981,23 km<sup>2</sup>, serta 33 kabupaten/kota yang dibagi menjadi 25 kabupaten dan 8 kota. Memiliki jumlah penduduk 15.115.206 jiwa dengan kepadatan penduduk 210 jiwa/km<sup>2</sup> pada tahun 2022.

Jumlah penduduk terus mengalami kenaikan dari tahun ke tahun. Ketika jumlah penduduk terus bertambah, berarti pemerintah juga harus terus menambah jumlah fasilitas hidup layak bagi masyarakatnya. Selain masalah tentang kependudukan, angka pengangguran di Provinsi Sumatera Utara juga masih tinggi. Menurut Badan Pusat Statistik, tingkat pengangguran terbuka dalam beberapa tahun terakhir meningkat secara signifikan. Angka tertinggi berada pada tahun 2022 mencapai 529.997 jiwa atau sekitar 6,91%. Dibandingkan tahun sebelumnya dengan tingkat pengangguran terbuka berjumlah 475.447 jiwa atau sekitar 6,33% [5]. Imbasnya, Jika strategi penciptaan lapangan kerja tidak diubah, pengangguran terbuka akan meningkat. Lihat gambar 1 di bawah ini untuk informasi lebih lanjut.



**Gambar 1.** Angka pengangguran terbuka dan jumlah penduduk Provinsi Sumatera Utara

Dari segi kesehatan, dalam evaluasi capaian pemerataan layanan kesehatan, cakupan kepesertaan Jaminan Kesehatan Nasional di Provinsi Sumatera Utara masih sebesar 72,4%, sehingga sasaran penguatan pelayanan kesehatan belum tercapai. Dalam sasaran Juga belum ada perbaikan dalam pencegahan dan pengendalian penyakit, atau penanganan kedaruratan kesehatan masyarakat, karena persebaran fasilitas kesehatan, dokter dan dokter spesialis yang tidak merata di Provinsi Sumatera Utara. Dan tujuan meningkatkan sumber daya untuk perawatan kesehatan juga tidak tercapai. Hal ini disebabkan oleh belum meratanya distribusi pelayanan kesehatan, terutama laboratorium cath yang terbatas, jumlah ahli jantung yang sedikit dan sulitnya menjangkau rumah sakit, meskipun akses transportasi sudah baik [6].

Hasil evaluasi Midterm Review Rencana Pembangunan Jangka Menengah Daerah (RPJMD) 2019-2023 (Pelaksanaan 2019-2020) dan hasil analisis data dan interpretasi informasi dari setiap persoalan pemerintahan

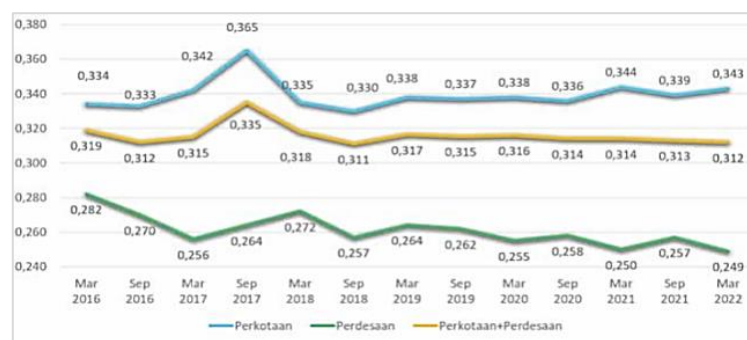
provinsi yang menjadi masalah terbesar yaitu dalam pembangunan daerah. Sumatera Utara yakni peningkatan kesejahteraan rakyat [7]. Maka dari itu dibutuhkan solusi guna mensukseskan hal tersebut. Yulianto menyatakan bahwa salah satu persyaratan keberhasilan program pembangunan sangat tergantung pada ketepatan identifikasi kelompok sasaran dan wilayah sasaran. Dengan mengidentifikasi karakteristik setiap daerah dengan cara mengelompokkan setiap daerah yang memiliki karakteristik yang sama, maka pemerintah dapat mengambil atau memutuskan kebijakan dan strategi yang baik/tepat untuk tujuan pembangunan [8].

Prioritas pembangunan Sumatera Utara telah dituangkan dalam dokumen Rencana Kerja Pemerintah Daerah (RKPD) yang merupakan implementasi dari RPJMD. Dokumen tersebut disusun dalam mengacu pada sasaran pembangunan, strategi dan prioritas pembangunan tahun 2019-2023. Sejarah operasional RKPD juga mengandung indikasi untuk meningkatkan efisiensi pemerintah dan perangkat daerah dalam melakukan tugas dan fungsi yang telah dibebankan kepadanya. Di sisi lain, secara fiksual merupakan acuan untuk menilai aktivitas dalam pelaksanaan kerja kemasyarakatan yang mewujudkan kesejahteraan rakyat.

Meningkatkan kesejahteraan masyarakat sejalan dengan adanya visi dan misi Provinsi Sumatera Utara. Visinya adalah membangun provinsi dengan jujur, benar, berani, tulus dan Ikhlas serta normatif. Visi ini berarti bahwa masyarakat berkomitmen untuk pembangunan yang lebih baik. Pembangunan terpadu dan berkelanjutan, serta pembangunan ekonomi dan manusia. Tujuan pembangunan ekonomi adalah untuk meningkatkan daya saing, berdaya saing melalui pemanfaatan sumber daya alam serta ilmu pengetahuan dan teknologi secara optimal. Tujuan pembangunan manusia adalah untuk menarik orang-orang yang berkompeten tinggi, jujur, dan religius.

Pemerintah Provinsi Sumatera Utara telah menyusun Strategi Kesejahteraan Manusia untuk meningkatkan kualitas, kuantitas dan kapasitas pendidikan, kesehatan serta dukungan sarana dan prasarana untuk kesejahteraan masyarakat. Tentunya untuk mencapai semua itu, Pemerintah Provinsi Sumatera Utara tidak bisa melakukannya sendiri. Diperlukan kerja kolaboratif antara negara, pemerintah daerah, perguruan tinggi, swasta dan masyarakat. Semuanya bermuara pada tujuan mewujudkan Provinsi yang maju, aman dan bermartabat.

Dari sisi kesejahteraan, pada bulan September 2022 jumlah penduduk miskin di Sumatera Utara berjumlah 1.262.090 jiwa, jumlah tersebut merupakan jumlah tertinggi dari semua provinsi di pulau Sumatera. Seiring bertambahnya jumlah orang miskin, demikian pula perbedaan pendapatan. Sejalan dengan kondisi tersebut, tingkat kedalaman dan keparahan kemiskinan juga meningkat. Dari segi ketimpangan, ketimpangan pendapatan di pedesaan relatif lebih berimbang dibandingkan di kota, sehingga ketimpangan pendapatan di kota harus diperhatikan. Ukuran ketimpangan yang umum digunakan adalah rasio gini. Nilai rasio gini bervariasi antara 0 dan 1 Semakin tinggi [9][10]. Secara grafik disajikan pada Gambar 2 berikut.



**Gambar 2.** Perkembangan Gini Ratio Sumatera Utara

Dalam Analisis kluster teknik yang digunakan untuk mengelompokkan suatu objek pengamatan sehingga setiap kluster menunjukkan kesamaan dengan subjek yang diteliti. Beberapa metode dapat digunakan untuk melakukan clustering, diantaranya adalah metode k-means clustering [11]. Dengan metode ini, sejumlah besar data dapat dikelompokkan dengan cepat dan efisien.

Berdasarkan penelitian terdahulu yang menggunakan K-Means Kluster diantaranya Rosyidi menggunakan metode K-means mengelompokkan 17 Kecamatan menjadi 3 kelompok kluster berdasarkan produksi bawang merah. Kluster 1 adalah kluster dengan produksi bawang merah rendah, kluster 2 adalah kluster dengan produksi bawang merah sedang, dan Kluster 3 adalah kluster dengan produksi bawang merah yang tinggi. Dimana karakteristik kluster yang terbentuk yaitu kluster pertama berisi kecamatan yang mempunyai jumlah luas tanam akhir, luas tambah tanam, luas panen, provitas, dan luas lahan sawah lebih sedikit dari rata-rata populasi. Kluster kedua berisi kecamatan yang mempunyai jumlah luas tanam akhir, luas tambah tanam, luas panen lebih sedikit dari rata-rata. Akan tetapi kluster ini mempunyai provitas dan luas lahan sawah lebih besar

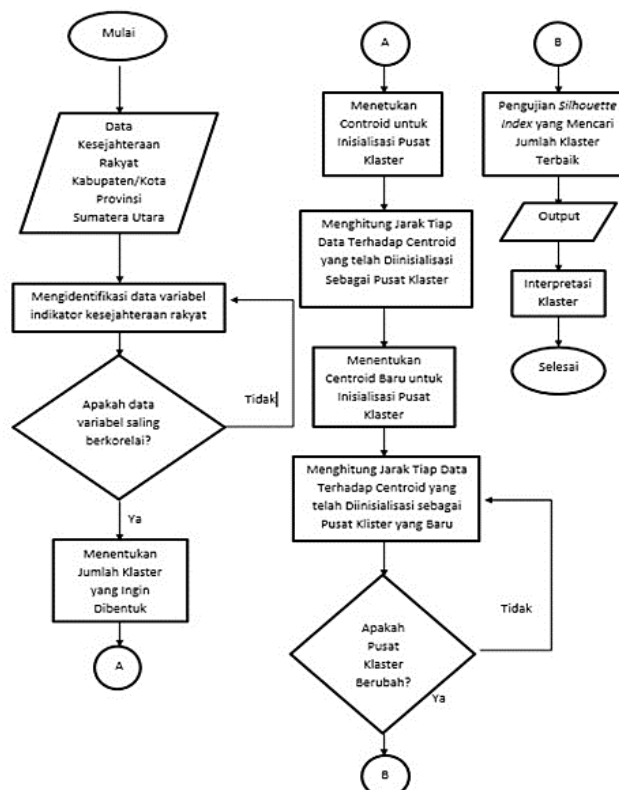
dari rata-rata populasi. Klaster ketiga berisi kecamatan yang mempunyai jumlah luas tanam akhir, luas tambah tanam, luas panen, provitas, dan luas lahan sawah lebih banyak dari rata-rata populasi [12].

Selain itu penelitian yang dilakukan oleh Anggi Dwi Lestari menggunakan metode K-means pengelompokan 35 Kabupaten/Kota menjadi 3 kelompok kluster berdasarkan Indeks Pembangunan Manusia (IPM). Klaster 1 adalah kelompok Kabupaten/Kota dengan IPM tinggi, klaster 2 adalah kelompok Kabupaten/Kota dengan IPM rendah, dan Cluster 3 adalah kelompok Kabupaten/Kota dengan IPM sedang. Variabel yang memberikan perbedaan paling besar pada ketiga cluster yang terbentuk adalah variabel harapan lama sekolah. Pemerintah seharusnya mengidentifikasi persoalan daerah dan membuat program-program yang mendukung capaian target IPM dan membangun IPM yang berkualitas tinggi [13].

Sedangkan penelitian yang dilakukan oleh Matius Tadi mengkaji tentang analisis kluster kemiskinan kabupaten/kota di Provinsi Banten dengan menggunakan metode K-Means diperoleh 2 kluster yaitu Klaster 1 yang terdiri dari Kota Tangerang, Kota Cilegon dan Kota Tangerang Selatan dan Klaster 2 yang terdiri dari Kecamatan Pandeglang, Kabupaten Lebak, Kabupaten Tangerang dan Kabupaten Serang. Kota Serang. Dan Cluster 2 memiliki skor *Structure Test* yang baik berdasarkan hasil *Silhouette Index*, sehingga Cluster 2 memiliki skor akurasi yang paling memiliki nilai keakuratan terbaik dibanding dengan klaster 3 dan klaster 4 [14]. Berdasarkan gambar latar belakang di atas, penulis tertarik untuk meneliti judul “Pengelompokan Kabupaten/Kota Provinsi Sumatera Utara Berdasarkan Indikator Kesejahteraan Rakyat Menggunakan K-Means Cluster”.

## 2. METODOLOGI PENELITIAN

Metode penelitian adalah sebagai cara ilmiah yang digunakan untuk mendapatkan data yang *objektif*, *valid*, dan *reliable* dengan tujuan dapat ditemukan, dibuktikan dan dikembangkan suatu pengetahuan [15]. Metode yang digunakan dalam penelitian ini adalah metode penelitian kuantitatif yang disajikan dalam bentuk deskriptif, menangkap ciri khas suatu objek, seseorang, atau suatu kejadian pada waktu data dikumpulkan dan ciri khas tersebut mungkin berubah dengan perkembangan waktu. Akan tetapi hal ini bukan berarti hasil penelitian waktu lalu tidak berguna, dari hasil-hasil tersebut kita dapat melihat perkembangan perubahan suatu fenomena dari masa ke masa.



Gambar 3. Alur Penelitian

### 2.1. Variabel Penelitian

Variabel merupakan semua yang ditetapkan oleh peneliti untuk dipelajari sehingga memperoleh informasi tentang hal tersebut kemudian ditarik kesimpulan [16]. Dalam penelitian ini berarti tidak ada variabel dependen ataupun independen karena analisis kluster termasuk pada kategori *interdependence techniques* [17].

Variabel-variabel yang digunakan dalam penelitian ini mengacu pada indikator-indikator kesejahteraan rakyat Provinsi Sumatera Utara [18], yaitu sebagai berikut.

**Tabel 1.** Variabel Kemiskinan Provinsi Sumatera Utara 2022

| No | Variabel                | Indikator                                |
|----|-------------------------|------------------------------------------|
| 1  | Kependudukan            | Jumlah penduduk (X1)                     |
|    |                         | Rasio Angka Beban Ketergantungan (X2)    |
|    |                         | Kepadatan Penduduk (X3)                  |
| 2  | Kesehatan               | Jumlah Tenaga Kesehatan (X4)             |
|    |                         | Fasilitas Kesehatan (X5)                 |
|    |                         | Imunisasi Campak (X6)                    |
| 3  | Pendidikan              | Angka Harapan Lama Sekolah (X7)          |
|    |                         | Rata-rata Lama Sekolah (X8)              |
|    |                         | Angka Partisipasi Murni SMA (X9)         |
| 4  | ketenagakerjaan         | Tingkat Partisipasi Angkatan Kerja (X10) |
|    |                         | Tingkat Pengangguran Terbuka (X11)       |
|    |                         | Angkatan Kerja (X12)                     |
| 5  | Taraf dan Pola Konsumsi | Konsumsi dan Pengeluaran (X13)           |
|    |                         | Rasio Gini (X14)                         |
| 6  | kemiskinan              | Persentase Penduduk Miskin (X15)         |
|    |                         | Indeks Kedalaman Kemiskinan (X16)        |
|    |                         | Indeks Keparahan Kemiskinan (X17)        |

## 2.2. Teknik Analisis Data

Analisis kluster adalah sebuah teknik yang digunakan untuk melakukan proses pengelompokan objek-objek pengamatan sehingga setiap group (kluster) memiliki kesamaan dalam hal tertentu yang sedang diteliti, dengan kata lain objek pengamatan dalam satu kluster dikatakan sama atau homogen. Terdapat perbedaan yang tinggi antar kluster dalam hal tertentu yang sedang diteliti, dengan kata lain objek pengamatan dalam kluster yang berbeda memiliki perbedaan yang nyata [19][20]. Analisis kluster (*cluster analysis*) merupakan suatu teknik statistika yang digunakan untuk mengelompokkan (grup) sekumpulan objek (manusia, produk, tanaman, dan sebagainya) ke dalam beberapa kluster [21][22][23].

Penelitian ini menggunakan analisis kluster dengan metode K-Means. Berikut adalah langkah-langkah pada penelitian ini: mendeskripsikan variabel dan informasi data indikator kesejahteraan rakyat di Provinsi Sumatera Utara tahun 2022. Melakukan uji asumsi multikolinieritas terhadap data variabel indikator kesejahteraan rakyat di Provinsi Sumatera Utara tahun 2022. Penelitian ini berasal dari nilai hubungan antar variabel apakah kurang dari 0.8 maka tidak terjadi multikolinieritas. Melakukan klusterisasi kabupaten/kota menggunakan metode K-Means. Dengan membentuk 3 percobaan kluster mulai 2 sampai 4 kluster. Melakukan evaluasi hasil kluster menggunakan *Silhouette Index* dari masing masing percobaan 2 sampai 4 kluster untuk mendapatkan kluster terbaik. Menginterpretasi hasil kluster.

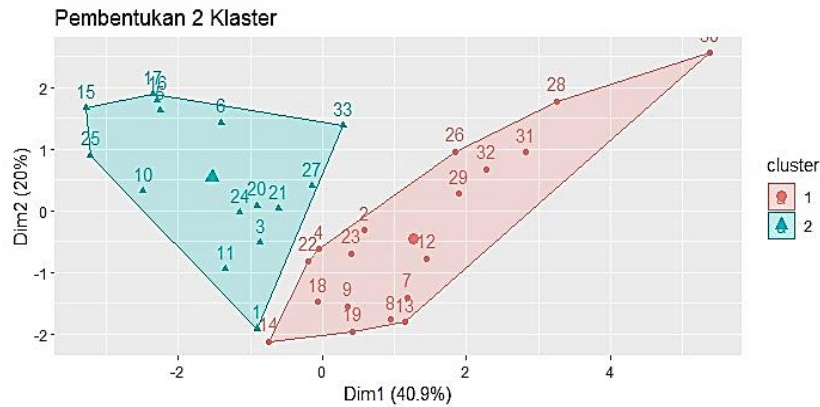
## 3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini berdasarkan ketersediaan data di Badan Pusat Statistik (BPS) Provinsi Sumatera Utara Tahun 2022. Karena isi data sangat bervariasi dalam satuan, dalam artian ada variabel (data) dengan satuan Jutaan (Jumlah penduduk), Puluhan (Kepadatan Penduduk), satuan belasan (Angka harapan lama sekolah) sampai satuan Jumlah di bawah 10 (Presentase Penududuk Miskin). Perbedaan yang sangat mencolok akan membuat perhitungan jarak (*distance*) menjadi tidak sehingga data asli harus di transformasi (standarisasi). Proses standarisasi menjadikan dua data dengan perbedaan satuan yang lebar akan otomatis menjadi menyempit. Standarisasi dilakukan karena metode k-means menggunakan konsep jarak antara objek/amatan, yang mana sensitif terhadap satuan pengukuran [17].

Sebelum dilakukan Pengelompokan Indikator Kesejahteraan Rakyat dengan menggunakan k-means, terlebih dahulu dilakukan uji multikolinieritas. Tetapi jika koefisien korelasi masing-masing variabel bebas  $< 0,8$  maka tidak terjadi multikolinearitas [24]. Jika koefisien korelasi masing-masing variabel bebas  $> 0,8$  maka terjadi multikolinearitas

Berdasarkan hasil pada di atas terlihat bahwa terdapat indikator yang memiliki nilai korelasi diatas 0.8 yaitu Indikator Jumlah penduduk ( $X_1$ ) dan Jumlah Tenaga kesehatan ( $X_4$ ) dengan nilai korelasi 0,88. Indikator Jumlah Penduduk ( $X_1$ ) dan Fasilitas Kesehatan ( $X_5$ ) dengan nilai korelasi 0,98. Indikator Jumlah Tenaga Kesehatan ( $X_4$ ) dan Fasilitas Kesehatan ( $X_5$ ) dengan nilai korelasi 0,92. Indikator Rata-rata Lama Sekolah ( $X_8$ ) dan Konsumsi dan Pengeluaran ( $X_{13}$ ) dengan nilai korelasi 0,81. Indikator Persentase Penduduk Miskin ( $X_{15}$ ) dan Indeks Kedalaman Kemiskinan ( $X_{16}$ ) dengan nilai korelasi 0,93. Indikator Persentase Penduduk Miskin ( $X_{15}$ ) dan Indeks Keparahan Kemiskinan ( $X_{16}$ ) dengan nilai korelasi 0,86. Indikator Indeks Kedalaman Kemiskinan ( $X_{16}$ ) dan Indeks Keparahan Kemiskinan ( $X_{17}$ ) dengan nilai korelasi 0,98. Dikarenakan ke-delapan Indikator





Gambar 6. Plot Pembentukan 2 Klaster

Dari gambar di atas pada pembentukan 2 klaster adalah sebagai berikut:

1. Klaster 1 terdiri dari Kabupaten Mandailing Natal, Tapanuli Tengah, Labuhan Batu, Asahan, Simalungun, Deli Serdang, Langkat, Nias Selatan, Serdang Bedagai, Batu Bara, Labuhanbatu Selatan, Labuanbatu Utara, Sibolga, Pematangsiantar, Tebing Tinggi, Medan, Binjai, Padangsidimpuan
2. Klaster 2 terdiri dari Kabupaten Nias, Tapanuli Selatan, Tapanuli Utara, Toba Samosir, Dairi, Karo, Humbang Hasundutan, Pakpak Bharat, Samosir, Padang Lawas Utara, Padang Lawas, Nias Utara, Nias Barat, Tanjungbalai, Gunungsitoli.

Tabel 3. Deskripsi Hasil Pola Pembentukan 2 Klaster

| Klaster | Nilai K | X2    | X3      | X6    | X7    | X9    | X10   | X11  | X12       | X14  |
|---------|---------|-------|---------|-------|-------|-------|-------|------|-----------|------|
| 1       | 18,17   | 48,63 | 1782,82 | 85,21 | 13,37 | 62,98 | 68,36 | 6,53 | 274317,20 | 0,28 |
| 2       | 15,60   | 56,36 | 229,16  | 70,70 | 13,33 | 78,38 | 79,30 | 2,40 | 627873,70 | 0,25 |

Klaster 1 pada kolom nilai K menunjukkan nilai 18,17 dari data kabupaten/kota yang telah diubah menjadi data numerik, yang artinya pusat klaster 1 diantara kabupaten mandailing. Pada variabel X2 dengan Nilai sebesar 48,63, X3 dengan Nilai sebesar 1782,82. X6 dengan niali sebesar 85,21, X7 dengan niali sebesar 13,37, X9 dengan nialisebesar 62,98, X10 dengan niali sebesar 68,36, X11 dengan niali sebesar 6,53, X12 dengan niali sebesar 273417,20, X14 dengan niali sebesar 0,28.

Klaster 2 pada kolom nilai K menunjukkan niali 15,60 dari data kabupaten/kota yang telah diubah menjadi data numerik. Yang artinya pusat kalster 2 dari antara kabupaten nias. Pada variabel X2 dengan nilai sebesar 56,36, X3 dengan nilai sebesar 229,16, X6 dengan nilai sebesar 70,70, X7 dengan niali sebesar 13,33, X9 dengan nilai sebesar 78,38, X10 dengan niali sebesar 79,30, X11 dengan nilai sebesar 2,40, X12 dengan niali sebesar 627873,70, X14 dengan niali sebesar 0,25.

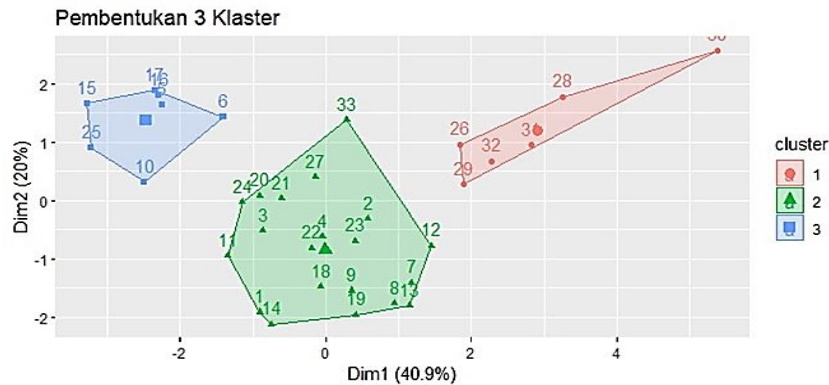
### 3.3. Pembentukan 3 Klaster

Berdasarkan hasil perhitungan tabel 4, setelah dilakukan pembentukan 3 klaster dengan pemilihan pusat klaster secara acak pada Rstudio. Diperoleh hasil sebagai berikut:

1. Klaster 1 berisi kabupaten/kota mempunyai rasio beban ketergantungan (X<sub>2</sub>), imunisasi campak (X<sub>6</sub>), angka partisipasi murni SMA (X<sub>9</sub>), dan tingkat pasrtisipasi angkatan kerja (X<sub>10</sub>) lebih kecil dari rata-rata keseluruhan. Sedangkan kepadatan penduduk (X<sub>3</sub>), angka harapan lama sekolah (X<sub>7</sub>) dan tingat pengangguran terbuka (X<sub>11</sub>), rasio gini (X<sub>14</sub>) lebih besar dari rata-rata keseluruhan. Akan tetapi angkatan kerja (X<sub>12</sub>) yang lebih sedikit dibanding rata-rata keseluruhan.
2. Klaster 2 berisi kabupaten/kota yang mempunyai rasio beban ketergantungan (X<sub>2</sub>), kepadatan penduduk (X<sub>3</sub>), angka harapan lama sekolah (X<sub>7</sub>), angka partisipasi murni SMA (X<sub>9</sub>), dan tingkat pasrtisipasi angkatan kerja (X<sub>10</sub>), rasio gini (X<sub>14</sub>) lebih kecil dari rata-rata keseluruhan. Sedangkan imunisasi campak (X<sub>6</sub>), dan tingat pengangguran terbuka (X<sub>11</sub>), lebih besar dari rata-rata keseluruhan. Akan tetapi angkatan kerja (X<sub>12</sub>) yang lebih banyak dibanding rata-rata keseluruhan.
3. Klaster 3 berisi kabupaten/kota yang mempunyai kepadatan penduduk (X<sub>3</sub>), tingat pengangguran terbuka (X<sub>11</sub>), dan rasio gini (X<sub>14</sub>) lebih kecil dari rata-rata keseluruhan. Sedangkan rasio beban ketergantungan (X<sub>2</sub>), imunisasi campak (X<sub>6</sub>), angka harapan lama sekolah (X<sub>7</sub>), angka partisipasi murni SMA (X<sub>9</sub>), tingkat pasrtisipasi angkatan kerja (X<sub>10</sub>), dan lebih besar dari rata-rata keseluruhan. Akan tetapi angkatan kerja (X<sub>12</sub>) yang lebih sedikit dibanding rata-rata keseluruhan.

**Tabel 4.** Hasil Perhitungan Pembentukan 3 Klaster

| klaster | X2    | X3    | X6    | X7    | X9    | X10   | X11   | X12   | X14   |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1       | -0,99 | 1,78  | -0,13 | 1,19  | -0,30 | -0,96 | 1,10  | -0,15 | 1,51  |
| 2       | -0,20 | -0,37 | 0,13  | -0,43 | -0,42 | -0,16 | 0,16  | 0,12  | -0,38 |
| 3       | 1,43  | -0,46 | -0,26 | 0,20  | 1,46  | 1,27  | -1,41 | -0,22 | -0,20 |

**Gambar 7.** Plot Pembentukan 3 Klaster

Dari gambar di atas dalam pembentukan 3 klaster adalah sebagai berikut:

1. Klaster 1 terdiri dari Kabupaten Tapanuli Utara, Toba Samosir, Dairi, Humbang Hasundutan, Pakpak Bharat, Samosir, Nias Barat.
2. Klaster 2 terdiri dari Kabupaten Nias, Mandailing Natal, Tapanuli Selatan, Tapanuli Tengah, Labuhan Batu, Asahan, Simalungun, Karo, Deli Serdang, Langkat, Nias Selatan, Serdang Bedagai, Batu Bara, Padang Lawas Utara, Padang Lawas, Labuhanbatu Selatan, Labuanbatu Utara, Nias Utara, Tanjungbalai, Gunungsitoli.
3. Klaster 3 terdiri dari Kabupaten Sibolga, Pematangsiantar, Tebing Tinggi, Medan, Binjai, Padangsidiempuan.

**Tabel 5.** Deskripsi Hasil Pola Pembentukan 3 Klaster

| klaster | Nilai K | X2    | X3      | X6    | X7    | X9    | X10   | X11  | X12       | X14  |
|---------|---------|-------|---------|-------|-------|-------|-------|------|-----------|------|
| 1       | 29,33   | 46,66 | 4824,63 | 76,57 | 14,00 | 65,86 | 65,61 | 7,64 | 252104,20 | 0,33 |
| 2       | 14,55   | 51,01 | 292,94  | 80,61 | 13,12 | 64,12 | 72,05 | 5,10 | 583378,80 | 0,25 |
| 3       | 13,43   | 60,06 | 103,07  | 74,66 | 13,46 | 90,25 | 83,62 | 0,82 | 167944,70 | 0,26 |

Klaster 1 pada kolom nilai K menunjukkan nilai 29,33 dari data kabupaten/kota yang telah dibuah menjadi data numerik, yang artinya pusat klaster 1 diantara kabupaten tapanuli utara. Pada variabel X2 dengan nilai sebesar 46,66, X3 dengan nilai sebesar 4824,6, X6 dengan nilai sebesar 76,57, X7 dengan nilai sebesar 14,00, X9 dengan nilai sebesar 65,86, X10 dengan nilai sebesar 65,61, X11 dengan nilai sebesar 7,64, X12 dengan nilai sebesar 252104,20, X14 dengan nilai sebesar 0,33.

Klaster 2 pada kolom nilai K menunjukkan nilai 14,55 dari data kabupaten/kota yang telah dibuah menjadi data numerik, yang artinya pusat klaster 2 diantara kabupaten nias. Pada variabel X2 dengan nilai sebesar 51,01, X3 dengan nilai sebesar 292,94, X6 dengan nilai sebesar 80,61, X7 dengan nilai sebesar 13,12, X9 dengan nilai sebesar 64,12, X10 dengan nilai sebesar 72,05, X11 dengan nilai sebesar 5,10, X12 dengan nilai sebesar 583378,80, X14 dengan nilai sebesar 0,25.

Klaster 3 pada kolom nilai K menunjukkan nilai 13,43 dari data kabupaten/kota yang telah dibuah menjadi data numerik, yang artinya pusat klaster 3 diantara kabupaten sibolga. Pada variabel X2 dengan nilai sebesar 60,06, X3 dengan nilai sebesar 103,07, X6 dengan nilai sebesar 74,66, X7 dengan nilai sebesar 13,46, X9 dengan nilai sebesar 90,25, X10 dengan nilai sebesar 83,62, X11 dengan nilai sebesar 0,82, X12 dengan nilai sebesar 167944,70, X14 dengan nilai sebesar 0,26.

### 3.4. Pembentukan 4 Klaster

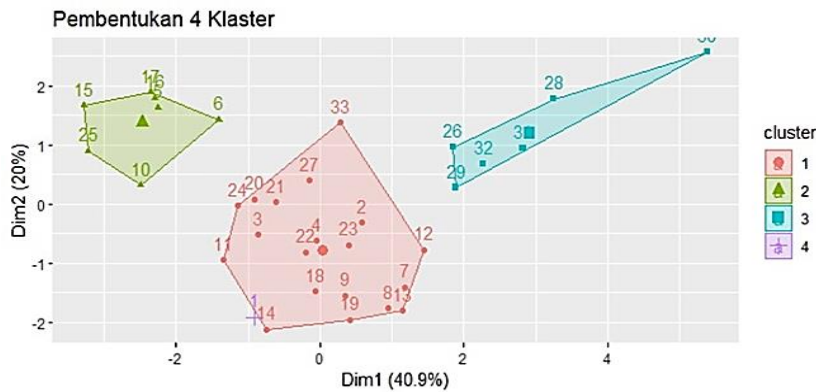
Berdasarkan hasil perhitungan tabel 6, setelah dilakukan pembentukan 4 klaster dengan pemilihan pusat klaster secara acak pada Rstudio. Diperoleh hasil sebagai berikut:

1. Klaster 1 berisi kabupaten/kota yang mempunyai rasio beban ketergantungan (X2), kepadatan penduduk (X3), angka harapan lama sekolah (X7), angka partisipasi murni SMA (X9), dan tingkat partisipasi angkatan kerja (X10) rasio gini (X14) lebih kecil dari rata-rata keseluruhan. Sedangkan imunisasi campak (X6) dan tingkat pengangguran terbuka (X11), lebih besar dari rata-rata

- keseluruhan. Akan tetapi angkatan kerja (X12) yang lebih sedikit dibanding rata-rata keseluruhan.
2. Klaster 2 berisi kabupaten/kota yang mempunyai kepadatan penduduk (X3), imunisasi campak (X6), tingkat pengangguran terbuka (X11), dan rasio gini (X14) lebih kecil dari rata-rata keseluruhan. Sedangkan rasio beban ketergantungan (X2), angka harapan lama sekolah (X7), angka partisipasi murni SMA (X9), Tingkat pasrtisipasi angkatan kerja (X10) dan lebih besar dari rata-rata keseluruhan. Akan tetapi angkatan kerja (X12) yang lebih sedikit dibanding rata-rata keseluruhan.
  3. Klaster 3 berisi kabupaten/kota yang mempunyai rasio beban ketergantungan (X2), imunisasi campak (X6), angka partisipasi murni SMA (X9), dan tingkat pasrtisipasi angkatan kerja (X10), lebih kecil dari rata-rata keseluruhan. Sedangkan kepadatan penduduk (X3), angka harapan lama sekolah (X7), tingkat pengangguran terbuka (X11), dan rasio gini (X14) lebih besar dari rata-rata keseluruhan. Akan tetapi angkatan kerja (X12) yang lebih sedikit dibanding rata-rata keseluruhan.
  4. Klaster 4 berisi kabupaten/kota yang mempunyai kepadatan penduduk (X3), angka harapan lama sekolah (X7), tingkat angka partisipasi murni SMA (X9), dan pengangguran terbuka (X11), lebih kecil dari rata-rata keseluruhan. Sedangkan rasio beban ketergantungan (X2), imunisasi campak (X6), tingkat pasrtisipasi angkatan kerja (X10), dan rasio gini (X14) lebih besar dari rata-rata keseluruhan. Akan tetapi angkatan kerja (X12) yang lebih banyak dibanding rata-rata keseluruhan.

**Tabel 6.** Hasil Perhitungan Pembentukan 3 Klaster

| klaster | X2    | X3    | X6    | X7    | X9    | X10   | X11   | X12   | X14   |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1       | -0,22 | -0,37 | 0,11  | -0,42 | -0,39 | -0,24 | 0,21  | -0,16 | -0,41 |
| 2       | 1,43  | -0,46 | -0,26 | 0,20  | 1,46  | 1,27  | -1,41 | -0,22 | -0,20 |
| 3       | -0,99 | 1,78  | -0,13 | 1,19  | -0,30 | -0,96 | 1,10  | -0,15 | 1,51  |
| 4       | 0,15  | -0,47 | 0,57  | -0,56 | -1,03 | 1,38  | -0,68 | 5,47  | 0,20  |



**Gambar 8.** Plot Pembentukan 4 Klaster

Dari gambar di atas dalam pembentukan 3 klaster adalah sebagai berikut:

1. Klaster 1 terdiri dari Kota Tapanuli Utara, Toba Samosir, Dairi, Humbang Hasundutan, Pakpak Bharat, Samosir, Nias Barat
2. Klaster 2 terdiri dari Kota Mandailing Natal, Tapanuli Selatan, Tapanuli Tengah, Labuhan Batu, Asahan, Simalungun, Karo, Deli Serdang, Langkat, Nias Selatan, Serdang Bedagai, Batu Bara, Padang Lawas Utara, Padang Lawas, Labuhanbatu Selatan, Labuanbatu Utara, Nias Utara, Tanjungbalai, Gunungsitoli.
3. Klaster 3 terdiri dari Kabupaten Sibolga, Pematangsiantar, Tebing Tinggi, Medan, Binjai, Padangsidempuan.
4. Klaster 4 terdiri dari Kota Nias.

**Tabel 7.** Deskripsi Hasil Pola Pembentukan 3 Klaster

| Klaster | K     | X2    | X3      | X6    | X7    | X9    | X10   | X11  | X12        | X14  |
|---------|-------|-------|---------|-------|-------|-------|-------|------|------------|------|
| 1       | 15,26 | 50,91 | 304,14  | 80,25 | 13,12 | 64,56 | 71,40 | 5,22 | 235273,80  | 0,25 |
| 2       | 13,43 | 60,06 | 103,07  | 74,66 | 13,46 | 90,25 | 83,62 | 0,82 | 167944,70  | 0,26 |
| 3       | 1,00  | 46,66 | 4824,63 | 76,57 | 14,00 | 65,86 | 65,61 | 7,64 | 252104,20  | 0,33 |
| 4       | 29,33 | 52,98 | 80,21   | 87,36 | 13,04 | 55,75 | 84,50 | 2,81 | 7197374,00 | 0,28 |

Klaster 1 pada kolom nilai K menunjukkan nilai 15,26 dari data kabupaten/kota yang telah dibuah menjadi data numerik, yang artinya pusat klaster 1 diantara kabupaten tapanuli utara. Pada variabel X2 dengan nilai sebesar 50,91, X3 dengan nilai sebesar 304,14, X6 dengan nilai sebesar 80,25, X7 dengan nilai sebesar 13,12, X9 dengan nilai sebesar 64,56, X10 dengan nilai sebesar 71,40, X11 dengan nilai sebesar 5,22, X12 dengan nilai 235273,80, X14 dengan nilai sebesar 0,25.

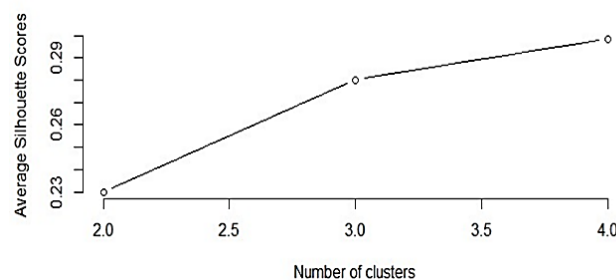
Klaster 2 pada kolom nilai K menunjukkan nilai 13,43 dari data kabupaten/kota yang telah dibuah menjadi data numerik, yang artinya pusat klaster 2 diantara kabupaten mandailing natal. Pada variabel X2 dengan nilai sebesar 60,06, X3 dengan nilai sebesar 103,07, X6 dengan nilai 74,66, X7 dengan nilai sebesar 13,46, X9 dengan nilai sebesar 90,25, X10 dengan nilai sebesar 83,62, X11 dengan nilai sebesar 0,82, X12 dengan nilai sebesar 252104,20, X14 dengan nilai sebesar 0,26.

Klaster 3 pada kolom nilai K menunjukkan nilai 1,00 dari data kabupaten/kota yang telah dibuah menjadi data numerik, yang artinya pusat klaster 3 diantara kabupaten sibolga. Pada variabel X2 dengan nilai sebesar 46,66, X3 dengan nilai sebesar 4824,63, X6 dengan nilai sebesar 76,57, X7 dengan nilai sebesar 14,00, X9 dengan nilai sebesar 65,86, X10 dengan nilai sebesar 65,61, X11 dengan nilai sebesar 7,64, X12 dengan nilai sebesar 252104,20, X14 dengan nilai sebesar 0,33.

Klaster 4 pada kolom nilai K menunjukkan nilai 29,33 dari data kabupaten/kota yang telah dibuah menjadi data numerik, yang artinya pusat klaster 4 diantara kabupaten nias. Pada variabel X2 dengan nilai sebesar 52,98, X3 dengan nilai sebesar 80,21, X6 dengan nilai sebesar 87,36, X7 dengan nilai sebesar 13,04, X9 dengan nilai sebesar 55,75, X10 dengan nilai sebesar 84,50, X11 dengan nilai sebesar 2,81, X12 dengan nilai sebesar 7197374,00, X14 dengan nilai sebesar 0,28.

### 3.5. Pengujian Klaster Terbaik

Mendeskripsikan variabel dan informasi data indikator kesejahteraan rakyat di Provinsi Sumatera Utara tahun 2022. Dengan melakukan uji asumsi multikolinieritas terhadap data variabel indikator kesejahteraan rakyat di Provinsi Sumatera Utara 2022, penelitian ini berasal dari nilai hubungan anatar variabel apakah  $< 0.8$  maka tidak terjadi multikolinieritas. Melakukan klasterisasi Kabupaten/Kota menggunakan metode K-Means . Dengan membentuk 3 percobaan klaster mulai dari 2 sampai 4 klaster. Melakukan evaluasi hasil kalster dengan menggunakan Silhouette Index dari masing-masing percobaan 2 sampai 4 klaster untuk medapatkan klaster terbaik. Seperti pada pada gambar dibawah ini.



**Gambar 9.** Hasil Pengujian Silhouette Index

Dalam gambar di atas menunjukkan hasil *silhouette index* bahwa pembentukan 4 klaster mendapat nilai tertinggi, yang artinya memiliki nilai keakuratan terbaik dibandingkan dengan pembentukan 2 klaster dan pembentukan 3 Klaster. Dalam menunjukkan struktur nilai yang lemah dari pembentukan klster dengan Nilai *silhouutte index* 0,2983784.

## 4. KESIMPULAN

Penelitian ini menggunakan analisis klaster metode K-Means . Penelitian diawali dengan standardisasi data asli karena isi data sangat bervariasi dalam satuan. Proses standardisasi menjadikan dua data dengan perbedaan satuan yang lebar akan otomatis menjadi menyempit. Transformasi data (standardisasi) dilakukan karena metode K-Means menggunakan konsep jarak antara objek/amatan, yang mana sensitif terhadap satuan pengukuran. Hasil dari pengujian Silhouette Index menunjukkan bahwa pembentuk 4 Klaster memiliki nilai keakuratan terbaik, berikutnya menentukan indikator kesejahteraan masyarakat pada masing-masing klaster. Berdasarkan tabel 4.4, didapat dua klaster yang terbentuk. Dengan menggunakan nilai rata-rata untuk setiap klaster maka dapat diketahui mana variabel yang dominan setiap klaster tersebut.

Kelebihan dari penelitian ini adalah dapat dijadikan acuan bagi pemerintah khususnya bagi pemerintah daerah dalam menentukan kebijakan-kebijakan yang tepat dalam merealisasikan peningkatan kesejahteraan masyarakat lebih teliti dan cermat dengan lebih memperhatikan faktor yang mempengaruhi di masing-masing daerah. Kelemahan dari penelitian ini adalah variabel kesejahteraan masyarakat tidak semua di catumkan dan belum mendalam.

**REFERENSI**

- [1] R. Raidani, "PENINGKATAN PENYEBARAN KESEJAHTERAAN RAKYAT UNTUK PENGELOMPOKAN KOTA MEDAN," *J. Badan Pengemb. Dan Penelit.*, vol. 10, no. 2, pp. 5–8, 2022.
- [2] A. Rahman, "Identifikasi Strategi Peningkatan Kesejahteraan Masyarakat di Kecamatan Sungaiambawang Kabupaten Kuburaya Provinsi Kalimantan Barat," *J. Manaj. Pembang.*, pp. 17–36, 2018.
- [3] E. Manita, M. Mawarpury, M. Khairani, and K. Sari, "Hubungan stres dan kesejahteraan (well-being) dengan moderasi kebersyukuran," *Gadjah Mada J. Psychol.*, vol. 5, no. 2, pp. 178–186, 2019.
- [4] L. V. Gorahe, F. Waani, and F. Tasik, "Dampak Pandemi Covid 19 Terhadap Kesejahteraan Masyarakat di Desa Dalako Bembanehe Kecamatan Tatoareng Kabupaten Kepulauan Sangihe," *J. Eksek.*, vol. 1, no. 1, 2021.
- [5] B. S. Utara, *Keadaan Ketenagakerjaan Agustus 2013*. Badan Pusat Statistik, 2013.
- [6] S. A. Nappoe and M. F. Kurniawan, "Evaluasi Capaian Pemerataan Layanan Kesehatan yang Berkeadilan di Era JKN di Provinsi Nusa Tenggara Timur".
- [7] P. D. S. Utara, *Rencana Pembangunan Jangka Menengah Daerah Sumatera Utara*. Sumatera Utara: Pemerintah Daerah Sumatera Utara, 2021. [Online]. Available: <https://sumutprov.go.id/artikel/halaman/rpjmd>
- [8] K. H. Hidayatullah and others, "Analisis klaster untuk pengelompokan kabupaten/kota di provinsi Jawa Tengah berdasarkan indikator kesejahteraan rakyat," *J. Stat. Univ. Muhammadiyah Semarang*, vol. 2, no. 1, 2014.
- [9] BPS, *Profil Kemiskinan Provinsi Sumatera Utara Maret 2023*, vol. 4. BPS Provinsi Sumatera Utara, 2023.
- [10] BPS, *Laporan Makro Sosial ekonomi Provinsi Sumatera Utara*. BPS Provinsi Sumatera Utara, 2023.
- [11] F. Gorunescu, *Data Mining: Concepts, models and techniques*, vol. 12. Springer Science & Business Media, 2011.
- [12] M. Rosyidi and I. Isnurani, "Analisis Klaster Metode K-Means Dalam Pengelompokan Kecamatan Di Kabupaten Brebes Berdasarkan Faktor-Faktor Produksi Bawang Merah," *Mathvision J. Mat.*, vol. 5, no. 1, pp. 16–21, 2023, [Online]. Available: <http://journal.unirow.ac.id/index.php/mv/issue/view/103>
- [13] Anggi Dwi Lestari, "Analisis Multivariat Clustering K-Means Pada Kabupaten/Kota di Provinsi Jawa Tengah Berdasarkan Indeks Pembangunan Manusia Tahun 2017 Dengan Bantuan Software SPSS," 2018, [Online]. Available: <http://lib.unnes.ac.id/36913/>
- [14] M. Tadi and B. A. Ningsi, "ANALISIS KLASSTER KEMISKINAN KABUPATEN KOTA DI PROVINSI BANTEN MENGGUNAKAN METODE K-MEANS," *J. Lebesque J. Ilm. Pendidik. Mat. Mat. dan Stat.*, vol. 4, no. 1, pp. 374–385, 2023.
- [15] R. Rasimin, "Metodologi Penelitian: Pendekatan Praktis Kualitatif." Mitra Cendikia, 2018.
- [16] Sugiyono, "Metode penelitian kuantitatif, kualitatif, dan R&D." Alfabeta, Bandung, 2017.
- [17] S. Santoso, "Statistik Multivariat Konsep dan Aplikasi dengan SPSS, Edisi revisi," *Jakarta PT. Elex Media Comput.*, 2014.
- [18] BPS, *Indikator Kesejahteraan Rakyat Provinsi Sumatera Utara 2022*. BPS Provinsi Sumatera Utara, 2022.
- [19] W. Hardle, L. Simar, and others, "Applied multivariate statistical analysis," 2003.
- [20] H. Jiawei and K. Micheline, *Data mining: concepts and techniques*. Morgan kaufmann, 2006.
- [21] P. U. Gio, "Belajar Statistika dengan R," 2018.
- [22] F. Murtagh and A. Heck, *Multivariate data analysis*, vol. 131. Springer Science & Business Media, 2012.
- [23] E. Prasetyo, "DATA MINING Mengolah Data Menjadi Informasi Menggunakan Matlab. In penerbit andi." 2014.
- [24] I. Ghozali, "Aplikasi analisis multivariete dengan program IBM SPSS 23," 2016.