

Artikel : [Akses terbuka/Open Access](#)

KLASIFIKASI KABUPATEN/KOTA DI PROVINSI NUSA TENGGARA TIMUR BERDASARKAN INDIKATOR STATUS KESEHATAN MASYARAKAT

Sitasi : Yunitaningtyas & Yolanda. 2022, JSTAR 2(1), 1-19.

Kronologi naskah.

Submit : 23 Mei 2021
Revisi : 13 November 2021
Diterima : 13 November 2021



Penyedia Data Statistik Berkualitas untuk
Indonesia Maju

REFORMASI BIROKRASI



KLASIFIKASI KABUPATEN/KOTA DI PROVINSI NUSA TENGGARA TIMUR BERDASARKAN INDIKATOR STATUS KESEHATAN MASYARAKAT*

Kristiana Yunitaningtyas¹, Anne Mudya Yolanda²

¹Pusat Kebijakan Pembiayaan dan Desentralisasi Kesehatan, Kementerian Kesehatan RI

²Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan
Alam, Universitas Riau, Indonesia

‡korespondensi author: kristianatyas@gmail.com, annemudyayolanda@lecturer.unri.ac.id.

Abstract

Segmented strategies are needed to improve the quality of public health. K-Means and K-Medoids clustering analysis could be applied to determine the distribution of public health indexes and to classify them by regencies/cities in Nusa Tenggara Timur province. This research conducted clustering analysis based on several public health indexes, namely Percentage of Women Aged 15-49 Years Who Have Given Birth in Last 2 Years with Medical Personnel Assistance, Percentage of Population Aged 0-59 Months by Complete Immunization, Percentage of Women Who Have Given Birth in the Last 2 Years by Normal Body Weight of Live Birth, Percentage of Households with Access to Sanitation, Percentage of Households with Access to Proper Drinking Water, and Percentage of Population Who Did not Smoke Tobacco in the Last Month. Based on Davies-Bouldin Index (DBI), it is known that K-Means is better than K-medoids Clustering. The First cluster of regencies/cities have best indicators, namely Timor Tengah Utara, Belu, Lembata, Flores Timur, Sikka, Ende, Ngada, Manggarai, Manggarai Barat, Malaka, Nagekeo, and Kupang City. The second cluster consists of Sumba Barat, Sumba Timur, and Sumba Tengah, have fairly good health indicators, but sanitation conditions need to be improved. The third cluster have fairly good health condition but government needs to pay attention about immunization in Alor, Kupang, Rote Ndao, Sumba Barat Daya, and Manggarai Timur. The fourth cluster, Timor Tengah Selatan and Sabu Raijua have low percentage of proper sanitation and complete immunization. Based on this, segmented health-related policies by public health current condition could be assigned.

Keyword: cluster analysis, k-means, k-medoids, public health index

1. Pendahuluan

Tercapainya kehidupan yang sehat dan terpenuhinya kesejahteraan bagi semua masyarakat tanpa memandang usia adalah salah satu tujuan dalam *Sustainable Development Goals* (SDGs)

atau Tujuan-Tujuan Pembangunan Berkelanjutan yang telah dirancang sebagai rencana aksi global hingga tahun 2030. Untuk mencapai kehidupan masyarakat yang sehat dan sejahtera maka diperlukan upaya pembangunan

kesehatan yang berkualitas. Kualitas pembangunan kesehatan ini dapat diukur hingga tingkat kabupaten/kota melalui Indikator Pembangunan Kesehatan Masyarakat (IPKM).

IPKM terdiri atas beberapa indikator kesehatan untuk merefleksikan masalah kesehatan masyarakat di suatu wilayah. Komposit indikator dalam IPKM adalah kesehatan ibu dan balita, penyakit menular, penyakit tidak menular, kesehatan reproduksi, dan status gizi. IPKM juga meliputi status determinan kesehatan yang mencakup aspek perilaku beresiko dan lingkungan.

Berdasarkan publikasi Badan Penelitian dan Pengembangan Kesehatan yang mengacu pada data hasil Riset Kesehatan Dasar (Riskesmas) Tahun 2018, nilai IPKM Provinsi Nusa Tenggara Timur Tahun 2018 sebesar 0.577. Nilai ini lebih tinggi dibandingkan dengan IPKM Tahun 2013 sebesar 0.462 (Kementerian Kesehatan RI, 2019). Secara umum terdapat peningkatan IPKM di Provinsi Nusa Tenggara Timur, namun unsur IPKM pada masing-masing kabupaten/kota menarik untuk diteliti lebih lanjut karena adanya dugaan perbedaan kualitas kesehatan antar kabupaten/kota. Dugaan ini berdasarkan data kesenjangan kabupaten/kota di Provinsi Nusa Tenggara Timur menurut IPKM 2018. Nilai sub indeks kesehatan balita kabupaten/kota di Nusa Tenggara Timur yang terendah adalah 0.4488 dan yang tertinggi adalah 0.7687. Untuk sub indeks pelayanan kesehatan di kabupaten/kota nilai yang paling rendah adalah 0.3065 dan yang tertinggi adalah

0.7325 serta nilai sub indeks kesehatan lingkungan yang terendah adalah 0.2164 dan yang tertinggi 0.8768. Adanya perbedaan kualitas ini dapat menggambarkan kondisi pembangunan kesehatan yang kurang merata.

Penelitian ini berfokus pada kondisi pembangunan kesehatan di 22 kabupaten/kota di Nusa Tenggara Timur berdasarkan IPKM dengan latar belakang perbedaan pendekatan oleh setiap daerah dalam meningkatkan status kesehatan masyarakat. Melalui IPKM maka dapat diperoleh peringkat kabupaten/kota dalam mengevaluasi pembangunan kesehatan dan sebagai dasar Penanggulangan Daerah Bermasalah Kesehatan (PDBK) (Kementerian Kesehatan RI, 2013). IPKM merupakan indikator penting untuk melihat kualitas penduduk selain Indeks Pembangunan Manusia (IPM) yang sudah lebih awam digunakan. Terdapat berbagai penelitian sebelumnya mengenai pengelompokan kabupaten/kota berdasarkan IPKM seperti pada kabupaten/kota di Jawa Timur (Putri & Fithriasari, 2015) namun belum terdapat penelitian terkait IPKM pada kabupaten/kota di Nusa Tenggara Timur. Penelitian yang dilakukan oleh Azkiyah (2017) telah mengelompokkan kabupaten/kota di Nusa Tenggara Timur namun berdasarkan Pelayanan Kesehatan Ibu dan Anak.

Salah satu metode yang dapat digunakan untuk mengetahui sebaran unsur-unsur IPKM menurut kabupaten/kota di wilayah Provinsi Nusa Tenggara Timur adalah *cluster*

analysis atau analisis gerombol yang menggabungkan objek-objek ke dalam beberapa kelompok berdasarkan karakteristik yang dimiliki. Pada analisis dilakukan perbandingan dua metode yaitu penggerombolan K-Means dan K-Medoids sehingga akan diperoleh metode yang lebih baik digunakan untuk mengelompokkan kabupaten/kota di Provinsi Nusa Tenggara Timur berdasarkan prevalensi enam unsur IPKM. Unsur-unsur IPKM yang diambil adalah persentase perempuan usia 15-49 tahun yang pernah melahirkan dalam 2 tahun terakhir dengan bantuan tenaga medis, persentase perempuan yang pernah melahirkan menurut berat badan anak lahir hidup, persentase penduduk usia 0-59 bulan menurut pemberian imunisasi lengkap, persentase rumah tangga dengan akses air minum layak, persentase rumah tangga dengan akses sanitasi layak, dan persentase penduduk yang dalam sebulan terakhir merokok tembakau.

Hasil yang diperoleh dalam penelitian ini dapat dimanfaatkan sebagai evaluasi sebaran capaian unsur-unsur indikator kesehatan masyarakat di masing-masing kabupaten/kota di Provinsi Nusa Tenggara Timur serta sebagai pendukung strategi berkelanjutan dalam peningkatan akses dan kualitas penyelenggaraan kesehatan. Peningkatan akses dan kualitas kesehatan masyarakat harus disesuaikan dengan kondisi daerah masing-masing. Kabupaten/kota dengan karakteristik unsur IPKM yang serupa perlu dikelompokkan agar kebutuhan

dari setiap klasifikasi kabupaten/kota dapat dipetakan. Selain itu, diharapkan upaya pembangunan kesehatan lebih efektif dan efisien untuk dilaksanakan.

2. Metodologi

Bahan dan Sumber Data

Penelitian ini menggunakan data sekunder dari publikasi Statistik Kesejahteraan Rakyat Nusa Tenggara Timur dan tabel/indikator sektor Badan Pusat Statistik (BPS) Nusa Tenggara Timur tahun 2019. Statistik Kesejahteraan Rakyat merupakan gabungan data-data statistik yang berkaitan dengan status kesejahteraan rakyat yaitu statistik kesehatan, statistik pendidikan, statistik perumahan serta statistik kependudukan, statistik fertilitas dan keluarga berencana, statistik teknologi informasi, dan statistik lainnya terkait kondisi sosial ekonomi (BPS, 2019). Publikasi Statistik Kesejahteraan Rakyat 2019 berasal dari pengolahan hasil Survei Sosial Ekonomi Nasional (Susenas) pada bulan Maret 2019 yang dirinci menurut kabupaten/kota di Provinsi Nusa Tenggara Timur. Penggunaan data pada Statistik Kesejahteraan Rakyat dikarenakan pada publikasi ini terdapat indikator-indikator yang menggambarkan kesejahteraan rakyat dari bidang kesehatan.

Terdapat 22 kabupaten/kota di Provinsi Nusa Tenggara Timur yang menjadi unit analisis dalam penelitian ini yaitu Kabupaten Sumba Barat, Kabupaten Sumba Timur, Kabupaten

Kupang, Kabupaten Timor Tengah Selatan, Kabupaten Timor Tengah Utara, Kabupaten Belu, Kabupaten Alor, Kabupaten Lembata, Kabupaten Flores Timur, Kabupaten Sikka, Kabupaten Ende, Kabupaten Ngada, Kabupaten Manggarai, Kabupaten Rote Ndao, Kabupaten Manggarai Barat, Kabupaten Sumba Tengah, Kabupaten Sumba Barat Daya, Kabupaten Nagekeo, Kabupaten Manggarai Timur, Kabupaten Sabu Raijua, Kabupaten Malaka, dan Kota Kupang.

Data indikator kesehatan yang digunakan terdiri atas enam variabel yang diukur pada setiap kabupaten/kota di Provinsi Nusa Tenggara Timur. Latar belakang pemilihan variabel persentase perempuan usia 15-49 tahun yang pernah melahirkan dalam 2 tahun terakhir dengan bantuan tenaga medis karena termasuk pada aspek pelayanan kesehatan, variabel persentase perempuan yang pernah melahirkan menurut berat badan anak lahir hidup dan persentase penduduk usia 0-59 bulan menurut pemberian imunisasi lengkap merupakan aspek kesehatan ibu dan anak, variabel persentase rumah tangga dengan akses air minum layak dan persentase rumah tangga dengan akses sanitasi layak termasuk dalam aspek kesehatan lingkungan, serta variabel persentase penduduk yang dalam sebulan terakhir merokok tembakau termasuk dalam aspek-aspek kesehatan lingkungan berdasarkan identifikasi faktor-faktor IPKM (Octavanny, Budiantara, & Ratnasari,

2017 dan Munawar, Aulia, Ferdhiana, Marzuki & Iqbal, 2020).

Masing-masing variabel dalam penelitian ini dijelaskan sebagai berikut:

1. Persentase Perempuan Usia 15-49 Tahun yang Pernah Melahirkan dalam 2 Tahun Terakhir dengan Bantuan Tenaga Medis.

Indikator ini merupakan perbandingan antara perempuan usia 15-49 tahun yang melakukan persalinan dibantu oleh tenaga medis dengan jumlah seluruh perempuan rentang usia tersebut yang pernah melahirkan dalam 2 tahun terakhir. Tenaga medis yang dimaksud adalah tenaga kesehatan seperti dokter, bidan, perawat, dan tenaga medis terlatih lainnya. Variabel ini sebagai indikator terkait Angka Kematian Ibu (AKI).

2. Persentase Perempuan yang Pernah Melahirkan Dalam 2 Tahun Terakhir Menurut Berat Badan Anak Lahir Hidup.

Indikator ini merupakan perbandingan antara perempuan yang melahirkan pada kategori berat badan anak lahir hidup kurang dari 2500 gram, sama dengan atau lebih dari 2500 gram, dan tidak ditimbang dengan jumlah seluruh perempuan yang pernah melahirkan anak lahir hidup dalam 2 tahun terakhir. Pada penelitian ini hanya digunakan persentase perempuan dalam 2 tahun terakhir yang melahirkan anak lahir hidup dengan berat badan sama dengan atau lebih dari 2500 gram sebagai indikator bahwa bayi yang

dilahirkan bukan bayi dengan berat badan lahir rendah (BBLR).

3. Persentase Penduduk Usia 0-59 Bulan Menurut Pemberian Imunisasi Lengkap.

Menurut Buku Ajar Imunisasi oleh Kementerian Kesehatan, pemberian imunisasi lengkap adalah balita yang mendapat imunisasi BCG sebanyak 1 kali, DPT sebanyak 3 kali, polio sebanyak 3 kali, campak 1 kali, dan hepatitis B sebanyak 3 kali. Indikator ini adalah perbandingan antara penduduk usia 0-59 bulan yang memperoleh imunisasi lengkap dengan seluruh penduduk usia 0-59 yang diimunisasi.

4. Persentase Rumah Tangga dengan Akses Air Minum Layak.

Menurut Sirusa BPS seperti yang dikutip oleh Mayasari (2019) akses air minum layak atau berkualitas adalah air minum yang terlindung meliputi ar ledeng (keran), keran umum, *hydrant* umum, terminal air, penampungan air hujan (PAH) atau mata air dan sumur terlindung, sumur bor atau sumur pompa, yang jaraknya minimal 10 meter dari pembuangan kotoran, penampungan limbah dan pembuangan sampah. Indikator ini adalah perbandingan antara rumah tangga yang dengan akses terhadap sumber air minum berkualitas (layak) dengan rumah tangga seluruhnya yang dinyatakan dalam persentase.

5. Persentase Rumah Tangga dengan Akses Sanitasi Layak.

Sanitasi layak adalah fasilitas sanitasi yang memenuhi syarat kesehatan

antara lain dengan kloset leher angsa atau plengsengan dengan tutup, tempat pembuangan akhir tinja menggunakan tanki septik atau Sistem Pengolahan Air Limbah (SPAL) dan fasilitas sanitasi tersebut digunakan oleh rumah tangga sendiri atau Bersama dengan rumah tangga lain tertentu menurut Sirusa BPS (Mayasari, 2019). Indikator ini adalah perbandingan antara rumah tangga dengan akses sanitasi yang layak atau berkualitas dengan rumah tangga seluruhnya.

6. Persentase Penduduk yang dalam Sebulan Terakhir Merokok Tembakau.

Indikator ini merefleksikan perbandingan penduduk dalam sebulan terakhir merokok tembakau yang dikategorikan setiap hari, tidak setiap hari, dan tidak merokok dengan seluruh jumlah penduduk. Pada penelitian ini hanya digunakan presentase penduduk dalam sebulan terakhir yang tidak merokok tembakau sebagai indikasi kesehatan masyarakat kabupaten/kota.

Keenam indikator ini merupakan variabel-variabel yang menjadi dasar pengelompokan 22 kabupaten/kota di Provinsi Nusa Tenggara Timur dari segi status kesehatan.

Metode Analisis Data

Data dianalisis menggunakan metode analisis gerombol dimana data kabupaten/kota dikelompokkan berdasarkan kedekatan sifat dari

masing-masing wilayah terkait dengan status kesehatan. Teknik analisis gerombol yang digunakan yaitu K-Means *Clustering* dan K-Medoids *Clustering*. Setelah dilakukan analisis maka akan dipilih metode yang lebih baik dalam mengelompokkan wilayah kabupaten/kota di Provinsi Nusa Tenggara Timur dengan membandingkan nilai Davies - Bouldin Index (DBI). Data dianalisis menggunakan perangkat lunak R.

Analisis Gerombol/Clustering

Pemodelan klasifikasi atau sering juga disebut metode klasifikasi adalah sebuah metode dari *data mining* yang digunakan untuk memprediksi kelompok suatu data pada kelompok kelas yang sudah ada sebelumnya (Izenman, 2008). Metode klasifikasi memungkinkan peneliti untuk mengklasifikasikan suatu objek baru ke dalam kelas tertentu berdasarkan nilai atribut-atributnya (Salazar, Velez & Salazar, 2012). Terdapat dua jenis metode klasifikasi yaitu *Unsupervised* dan *Supervised*.

Metode klasifikasi *supervised* sering digunakan untuk membuat klasifikasi pada data yang telah memiliki kelompok atau label tertentu. Pada pendekatan *supervised*, suatu data baru akan dikelompokkan berdasarkan data *training* setelah diketahui informasi atau kelas antar grupnya. Metode ini akan menghasilkan aturan pembeda atau batas antar kelompok. Aturan tersebut nantinya dapat digunakan untuk

mengklasifikasikan kelompok dari amatan lain yang belum tercakup dalam data.

Metode klasifikasi *unsupervised* digunakan untuk mengklasifikasikan kelompok yang mana sebelumnya tidak terdapat informasi mengenai kelompok dari amatan. Teknik ini biasa digunakan untuk menggambarkan struktur tersembunyi yang terdapat pada data tanpa adanya informasi atau label mengenai kelompok atau grup dari amatan pada data yang digunakan. Metode ini mengelompokkan data atau dengan kata lain menentukan keanggotaan grup dari amatan berdasarkan persamaan tertentu dan mengurangi dimensi data, tetapi tetap mempertahankan struktur dan kegunaannya. Analisis *unsupervised* ini sering disebut sebagai analisis gerombol (*Clustering* atau *Cluster Analysis*).

Analisis gerombol atau analisis *clustering* merupakan teknik untuk mengidentifikasi kelompok dari suatu observasi sehingga dapat diperoleh informasi mengenai karakteristiknya. Analisis ini merupakan proses untuk mengorganisir data ke dalam kategori tertentu sehingga selanjutnya dapat digunakan secara lebih efektif dan efisien. Algoritma yang dijalankan dalam analisis gerombol yaitu mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam *cluster* yang sama (Debataraja, 2019).

Analisis gerombol umumnya digunakan untuk analisis eksplorasi pada

amatan dan tidak ada pemodelan atau ada prediksi yang dibuat (EMC Education Services, 2015). Tujuan utama dalam analisis gerombol yaitu mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Karakteristik tersebut dapat berupa variabel-variabel tertentu yang menjadi indikator dari amatan.

Secara umum, tahapan dalam analisis gerombol mulai dari mendapatkan sumber data dengan karakteristik tertentu sampai evaluasi hasil pengelompokan (penggerombolan). Tahapan tersebut mencakup tiga bagian utama, yaitu pemilihan fitur/transformasi, pemilihan algoritma atau teknik pengelompokan (penggerombolan), dan diakhiri dengan melakukan evaluasi hasil pengelompokan (Liu, Ren, & Cai, 2020). Dalam menyelesaikan pengelompokan tersebut ada tiga hal utama yang harus dilakukan, yaitu mencari ukuran kemiripan antar amatan, cara membentuk kelompok atau gerombol, dan menentukan banyak kelompok atau gerombol yang akan dibuat. Oleh karenanya, diperlukan suatu teknik perhitungan untuk dapat menjawab tiga hal dasar ini, diantaranya Algoritma *Average Silhouette* dan *Gap Statistics*.

Average Silhouette

Algoritma *Average Silhouette* yang diperkenalkan oleh Peter J. Rousseuw adalah alat pengukuran untuk menilai validitas dari penggerombolan.

Silhouette dibentuk sehingga diperoleh jumlah gerombol optimal dengan data berskala rasio seperti jarak Euclidean yang cocok untuk memisahkan antar kelompok dengan jelas (Thinsungnoen, Kaoungku, Durongdumronchai, Kerdprasop, & Kerdprasop, 2015). Dasar dari penentuan nilai *silhouette* yaitu pengukuran kemiripan suatu objek dengan kelompoknya dibandingkan dengan kelompok lain. Persamaan dari *silhouette* dapat dijelaskan sebagai berikut:

$$S_k = (b_k - a_k) / \max(a_k, b_k)$$

Untuk tiap k dihitung a_k yang merupakan rata-rata jarak dari k dengan anggota lain pada gerombol yang sama. Untuk tiap k dan masing-masing gerombol k yang bukan anggota gerombol j dihitung rata-rata jarak dari j dengan anggota gerombol k yaitu b_{kj} .

Pemecahan dari setiap kelompok induk diperoleh rata-rata yang disebut sebagai *split silhouette* (SS_j) dengan $j = 1, 2, \dots, c$. *Split silhouette* merupakan ukuran dari gerombol yang heterogen (Wulandari, 2020).

Gap Statistics

Algoritma *Gap Statistics* adalah salah satu metode yang digunakan untuk menentukan banyaknya *clusters* atau gerombol dari data tanpa diketahui banyaknya klasifikasi. Algoritma ini diperkenalkan oleh Tibshirani (Tibshirani, Walther, & Hastie, 2001) dan diperoleh melalui metode sampling Monte Carlo. *Gap Statistics* menghitung

jumlah kuadrat dari jarak Euclidean diantara dua pengukuran pada masing-masing kelas. Hasil penggerombolan dari distribusi zero-mean yang dibentuk lalu dibandingkan untuk memperoleh banyak gerombol yang optimal dari data (Yuan & Yang, 2019).

Gap Statistics untuk besarnya k tertentu yaitu:

$$Gap_n(k) = E_n^* \log(W_k) - \log(W_k)$$

dengan E_n^* merupakan ukuran sampel n dan $\log(W_k)$ merupakan log rata-rata. Langkah-langkah dalam penentuan jumlah gerombol yang optimal dengan *Gap Statistics* sebagai berikut (Dewa & Jatipaningrum, 2019):

1. Dataset yang akan diteliti digabungkan dan dicoba variasi jumlah gerombol dari $k = 1, \dots, k_{max}$. Kemudian dihitung total yang sesuai dalam variasi intra-cluster W_k .
2. Dibentuk data referensi dengan membangkitkan data berdasarkan distribusi acak seragam. Dilakukan penggerombolan dari data referensi ini dengan berbagai jumlah gerombol $k = 1, \dots, k_{max}$ dan dihitung total yang sesuai dalam variasi intra-cluster W_{kb} .
3. Estimasi *Gap Statistics* dihitung sebagai penyimpangan dari nilai W_k yang diamati dari W_{kb} dan dihitung juga nilai standar deviasinya.
4. Dipilih jumlah gerombol sebagai nilai terkecil dari k sehingga nilai gap statistics berada dalam satu

standar deviasi dari $k + 1$.

K-Means Clustering

K-Means merupakan teknik untuk pengklasifikasian karakteristik dari objek. K-Means adalah salah satu bentuk analisis gerombol atau *clustering* yang paling sederhana dan sangat luas digunakan pada data tanpa label tertentu (*unlabeled data*). Konsep dari algoritma K-Means memisahkan data menjadi beberapa bagian yang terpisah.

Pengelompokan dengan K-Means dimulai secara acak dan dilakukan berulang hingga ditemukan kelompok dengan karakteristik yang sama. Amatan yang memiliki karakteristik yang sama atau dalam hal ini dikatakan memiliki kemiripan akan dikelompokkan dalam satu kelompok yang sama. Amatan lain dengan karakteristik yang berbeda dikelompokkan dengan kelompok yang lain sehingga data yang berada dalam satu kelompok memiliki tingkat variasi yang kecil (Kantardzic, 2011).

K-Means menggunakan jarak untuk membagi data numerik ke dalam sejumlah kelompok. Luaran dari analisis K-Means adalah melakukan klasifikasi data dengan memaksimalkan kemiripan data dalam satu gerombol dan meminimalkan kemiripan data antar gerombol. Ukuran kemiripan yang digunakan dalam cluster adalah fungsi jarak. K-Means *Clustering* akan mengklasifikasikan data set tersebut menjadi sebanyak k gerombol dengan masing-masing gerombol memiliki ciri

tertentu berdasarkan titik pusat gerombol (*centroid*) terdekat dengan data. Oleh karena itu, pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data terhadap titik *centroid* (Thrun, 2017).

K-Medoids Clustering

Suatu data set dapat berisi amatan yang sangat besar atau sangat kecil dibandingkan kumpulan data lainnya, analisis K-means sangat rentan terhadap efek pencilan. Metode ini mengasumsikan semua klaster memiliki jumlah pengamatan yang sama yang mungkin tidak selalu terjadi terutama jika ada lebih banyak pencilan dengan nilai ekstrim. Pemilihan *centroid* dilakukan secara acak dalam metode ini, maka dapat menghasilkan *centroid* yang berbeda bila dilakukan berkali-kali bahkan dalam kondisi yang sama membatasi pengulangan hasil. Masalah ini dapat diatasi dengan melakukan teknik pengelompokan lain yang disebut metode K-Medoids (Soni & Patel, 2017).

Teknik penggerombolan K-Medoids atau juga disebut *Partitioning Around Medoid* (PAM) merupakan alternatif dari algoritma K-Means dan termasuk dalam algoritma *partitioning*. Pada algoritma *partitioning*, data dikelompokkan ke dalam sejumlah gerombol tanpa adanya struktur hirarki. Analisis K-Medoids secara umum bekerja dengan meminimalkan jumlah jarak dari titik amatan dalam sebuah gerombol ke satu titik data yang dipilih

dalam gerombol yang sama, yang disebut medoid. Jarak rata-rata antara medoid dan subset titik data dalam klaster yang sama diminimalkan (Thrun, 2017). Algoritma K-Medoids digunakan untuk menemukan Medoids dalam sebuah klaster yang merupakan titik pusat cluster (*centroid*). K-Medoids lebih *robust* dibandingkan dengan K-Means. Analisis K-Medoids mampu bekerja pada amatan dengan *noise* dan pencilan (Arora, Deepali, & Varshney, 2016).

Algoritma K-Medoids dan K-Means sebenarnya serupa. Selain perubahan dari *centroid* ke *medoid*, algoritma K-Means dapat diformulasikan secara analog ke K-Medoids. Tahapan dalam algoritma K-Medoids dimulai dari menentukan nilai k sebagai medoid awal. Proses dilanjutkan dengan menghitung jarak objek data ke pusat gerombol secara acak dari n objek data sedemikian rupa sehingga partisi awal dibuat berdasarkan kedekatan masing-masing objek ke *centroid* untuk memulai proses partisi data. Kemudian metode iterasi digunakan secara terus menerus hingga diperoleh nilai partisi yang paling sesuai (Kumar & Sirohi, 2010).

Davies - Bouldin Index

Davies-Bouldin Index merupakan salah satu metode yang digunakan untuk mengukur validitas dari gerombol dari suatu metode penggerombolan. Metode Davies – Bouldin Index diusulkan oleh David L. Davies dan

Donald W. Bouldin pada tahun 1979. Pengukuran menggunakan Davies Bouldin Index memaksimalkan jarak antar gerombol dan pada saat bersamaan mencoba untuk meminimalkan jarak antar objek dalam gerombol. Apabila jarak antar gerombol maksimum berarti terdapat sedikit persamaan antara masing-masing gerombol sehingga perbedaan antar gerombol terlihat lebih jelas. Jika jarak antar gerombol minimal, maka setiap objek dalam gerombol memiliki persamaan karakteristik yang tinggi.

Hasil penggerombolan diperoleh dari penentuan titik tengah (*centroid*) yang kemudian dievaluasi dengan Davies - Bouldin Index sehingga dapat diketahui korelasi dari metode penentuan *centroid* berdasarkan *sum of squared error* dengan kualitas gerombol berdasarkan nilai Davies - Bouldin Index yang diperoleh (Jumadi Dehotman Sitompul, Salim Sitompul, & Sihombing, 2019). Persamaan dari Davies - Bouldin Index adalah (Hardiani, Sulistyio, & Hartanto, 2015):

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} \left\{ \frac{\Delta(i) + \Delta(j)}{D(i, j)} \right\}$$

dengan $D(i, j) = \min_{x \in S_i, y \in S_j} dist(x, y)$

$$\Delta(i) = \max_{x, y \in S_i} dist(x, y)$$

$$\Delta(j) = \max_{x, y \in S_j} dist(x, y)$$

Nilai Davies - Bouldin Index yang kecil merupakan jumlah gerombol yang baik. Semakin kecil nilai Davies - Bouldin Index maka semakin optimal hasil gerombol.

3. Hasil dan Pembahasan

Eksplorasi Data

Statistik Kesejahteraan Rakyat dari Bidang Kesehatan Tahun 2019 di Provinsi Nusa Tenggara Timur secara umum cukup baik. Persentase perempuan usia 15-49 tahun yang pernah melahirkan dalam 2 tahun terakhir dengan bantuan tenaga medis untuk tingkat provinsi mencapai angka 83.19, persentase perempuan yang pernah melahirkan dalam 2 tahun terakhir menurut berat badan anak lahir hidup untuk berat badan lahir normal sebesar 73.0. Angka yang cukup rendah adalah persentase penduduk usia 0-59 bulan menurut pemberian imunisasi lengkap untuk tingkat provinsi sebesar 51.59.

Pada indikator air layak minum, statistik menunjukkan angka yang bagus, yaitu persentase rumah tangga dengan akses air minum layak untuk skala provinsi adalah sebesar 82.35. Pada sisi lain, persentase rumah tangga dengan akses sanitasi layak Provinsi Nusa Tenggara Timur masih berada di bawah 70 persen yaitu hanya mencapai angka 64.55 persen. Angka-angka ini tentunya masih perlu ditingkatkan, terutama agar terjadi pemerataan di wilayah kabupaten/kota.

Indikator terakhir yaitu persentase penduduk yang dalam sebulan terakhir tidak merokok tembakau menunjukkan angka yang baik. Pada tahun 2019 diperoleh persentase penduduk yang

dalam sebulan terakhir tidak merokok tembakau yaitu 79.71. Keadaan ini tentunya cukup baik, mengingat merokok berkaitan erat dengan Kondisi kesehatan dalam hal menimbulkan atau memperparah beberapa jenis penyakit. Indikator kesehatan untuk kabupaten/kota maupun Provinsi Nusa Tenggara Timur ditampilkan pada Tabel 1.

Tabel 1. Indikator Kesehatan (%) Kabupaten/Kota di Provinsi Nusa Tenggara Timur Tahun 2019.

Wilayah	X1	X2	X3	X4	X5	X6
Sumba Barat	92.55	38.21	75.6	37.70	70.41	77.11
Sumba Timur	88.74	61.88	74.04	49.62	67.35	73.45
Kupang	72.47	44.4	69.41	67.15	79.66	81.39
Timor Tengah Selatan	78.07	34.2	68.1	35.03	55.5	86.73
Timor Tengah Utara	91.07	68.34	77.93	62.67	83.12	78.74
Belu	91.82	51.18	73.97	64.9	85.65	81
Alor	68.7	33.99	70.18	83.9	79	77.56
Lembata	100	56.04	88.37	90.36	93.6	81.51
Flores Timur	97.24	66.85	87.94	86.38	84.81	78.17
Sikka	99.12	63.22	87.22	79.21	88.67	75.27
Ende	89.65	58.25	73.77	91.49	94.17	80.86
Ngada	99.04	51.44	86.7	82.96	95.16	77.58
Manggarai	94.87	67.44	85.75	56.14	91.26	77.99
Rote Ndao	61.26	45.18	51.09	74.85	87.61	80.29
Manggarai Barat	86.24	65.16	80.15	52.27	85.6	76.27
Sumba Tengah	87.49	48.77	78.06	31.8	55.35	73.95

Sumba Barat Daya	53.89	28.46	47.35	28.24	67.38	82.28
Nagekeo	100	64.81	83.85	81.61	96.7	79.78
Manggarai Timur	64.52	46.3	53.49	36.86	84.02	75.95
Sabu Raijua	75.21	42.4	68.35	78.62	45.3	85.48
Malaka	97.32	54.88	62.66	67.28	89.32	83.18
Kota Kupang	89.82	59.8	86.84	88.02	98.95	81.56
Provinsi NTT	83.19	73.0	51.59	82.35	64.55	79.71

Sumber data: BPS Provinsi NTT

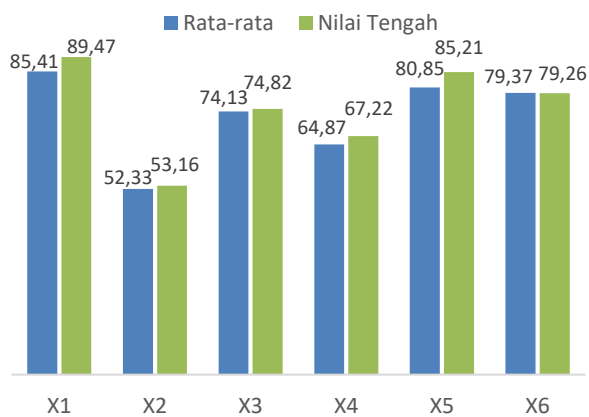
Keterangan:

- X1= Persentase Perempuan Usia 15-49 Tahun yang Pernah Melahirkan dalam 2 Tahun Terakhir dengan Bantuan Tenaga Medis (Persalinan Medis)
- X2= Persentase Penduduk Usia 0-59 Bulan Menurut Pemberian Imunisasi Lengkap (Imunisasi Lengkap)
- X3= Persentase Perempuan yang Pernah Melahirkan Dalam 2 Tahun Terakhir Menurut Berat Badan Anak Lahir Hidup dengan badan lahir normal (BBL Normal)
- X4= Persentase Rumah Tangga dengan Akses Sanitasi Layak (Sanitasi Layak)
- X5= Persentase Rumah Tangga dengan Akses Air Minum Layak (Air Minum Layak)
- X6= Persentase Penduduk yang dalam Sebulan Terakhir Tidak Merokok Tembakau (Tidak Merokok)

Berdasarkan data yang disajikan pada Tabel 1 dapat dihitung rata-rata persentase dari masing-masing variabel indikator yang ditampilkan pada Gambar 1.

Persentase perempuan usia 15-49 tahun yang pernah melahirkan dalam 2 tahun terakhir dengan bantuan tenaga medis (persalinan medis) diperoleh rata-rata sebesar 85.41 persen dan nilai tengah

sebesar 89.47 persen. Angka ini menunjukkan secara rata-rata persentase di kabupaten/kota berada di atas persentase provinsi. Terdapat daerah yang nilainya berada di bawah persentase provinsi berdasarkan Tabel 1 yaitu Kabupaten Sumba Barat Daya (53.89 persen) namun ada daerah dengan nilai sempurna (100 persen) yaitu Kabupaten Lembata dan Kabupaten Nagekeo.



Gambar 1. Rata-rata dan Nilai Tengah Indikator Kesehatan (%) Kabupaten/Kota di Provinsi Nusa Tenggara Timur Tahun 2019.

Pada variabel imunisasi lengkap, rata-rata dan nilai tengah persentase penduduk usia 0-59 bulan yang mendapat imunisasi lengkap masing-masing adalah 52.33 persen dan 53.16 persen. Berdasarkan Tabel 1 persentase terbesar dicapai oleh Kabupaten Timor Tengah Utara (68.34 persen) dan wilayah dengan persentase terendah adalah Kabupaten Sumba Barat Daya (28.46 persen).

Wilayah dengan persentase perempuan yang pernah melahirkan dalam 2 tahun terakhir menurut berat

badan anak lahir hidup dengan badan lahir normal (BBL Normal) paling rendah adalah Kabupaten Sumba Barat Daya sebesar 47.35 persen menurut Tabel 1. Angka ini memiliki rentang cukup jauh dengan persentase terbesar yaitu 88.37 persen dicapai oleh Kabupaten Lembata. Berdasarkan hasil perhitungan diketahui rata-rata dan nilai tengah dari variabel BBL normal sebesar 74.13 persen dan 74.82 persen.

Adapun untuk rumah tangga dengan akses sanitasi layak menurut Tabel 1, persentase paling besar yaitu 91.49 persen dicapai oleh Kabupaten Ende dan paling rendah Kabupaten Sumba Barat Daya sebesar 28.24 persen. Nilai ini juga memiliki jangkauan cukup jauh (63.25), yang mana diketahui rata-rata dan nilai tengah masing-masing sebesar 64.87 persen dan 67.22 persen.

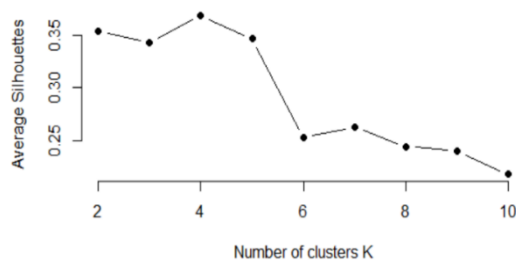
Variabel air minum layak memiliki nilai rata-rata sebesar 80.85 persen dan nilai tengah sebesar 85.21 persen yang sangat baik. Bahkan berdasarkan Tabel 1 persentase rumah tangga dengan akses air minum layak (air minum layak) paling tinggi mencapai angka 98.95 persen (Kota Kupang). Hal ini dapat didukung oleh posisi Kota Kupang sebagai ibu kota provinsi. Kebijakan terkait air minum layak tentunya masih harus dirumuskan, karena masih terdapat ketimpangan, yang mana ada daerah yang persentase air minum layaknya masih berada pada level 45.3 persen yaitu Kabupaten Sabu Raijua.

Statistik berupa rata-rata dan nilai tengah dari Penduduk yang dalam

sebulan terakhir tidak merokok tembakau adalah 79.37 persen dan 79.26 persen. Perbedaan nilai pada variabel ini cukup kecil jika dilihat dari jangkauan antar data. Artinya, persentase antar kabupaten/kota tidak berbeda signifikan. Menurut Tabel 1 persentase tertinggi yaitu 86.73 persen untuk Kabupaten Timor Tengah Selatan dan terendah sebesar 73.45 persen pada Kabupaten Sumba Timur.

Penentuan k optimal

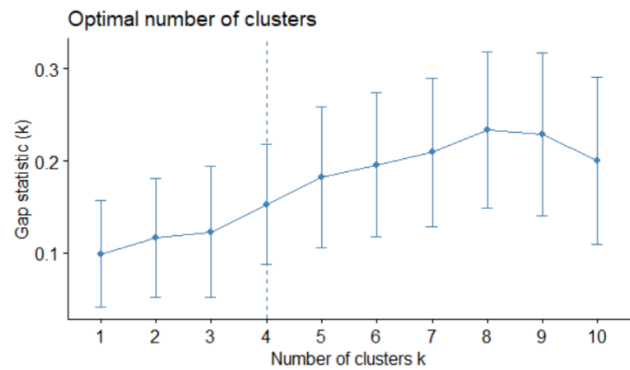
Penentuan besarnya nilai k yang optimal bagi data menggunakan metode *Average Silhouette* dan *Gap Statistics*. Menggunakan *Average Silhouette* diperoleh grafik pada Gambar 2. Berdasarkan hasil dari algoritma *Average Silhouette*, banyak gerombol yang optimal digunakan untuk mengelompokkan data pada penelitian ini yaitu 4 gerombol. Ini terlihat dari Gambar 2 dimana nilai *Average Silhouette* yang tertinggi diperoleh saat k bernilai 4.



Gambar 2. Banyaknya k atau gerombol yang optimal dengan metode *Average Silhouette*.

Untuk memperoleh nilai k atau banyaknya gerombol yang optimal juga dapat digunakan metode *Gap Statistics*. Hasil penentuan banyak gerombol

menggunakan metode ini ditampilkan pada Gambar 3.

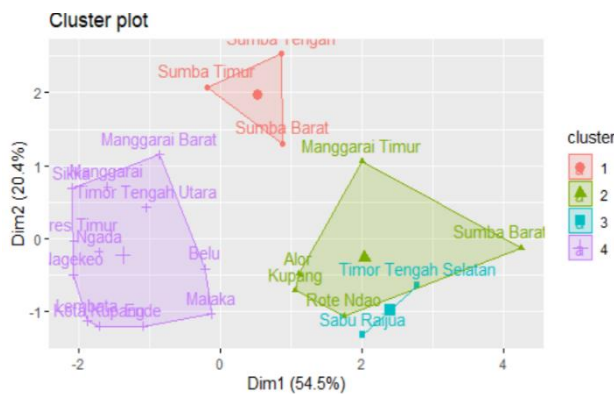


Gambar 3. Banyaknya k atau gerombol yang optimal dengan metode *Gap Statistics*.

Menggunakan metode *Gap Statistics* diperoleh nilai k atau banyak gerombol yang optimal adalah 4 gerombol. Jadi berdasarkan metode *Average Silhouette* maupun *Gap Statistics* nilai k atau banyaknya gerombol yang optimal digunakan untuk klasifikasi kabupaten/kota di Provinsi Nusa Tenggara Timur adalah sebanyak 4 gerombol. Selanjutnya dilakukan analisis gerombol dengan teknik K-Means dan K-Medoids menggunakan k atau gerombol sebanyak 4.

Analisis Gerombol dengan K-Means dan K-Medoids

Teknik analisis gerombol K-Means yang diterapkan pada data berdasarkan banyaknya k optimal yaitu $k=4$ menghasilkan kelompok-kelompok seperti pada Gambar 4.



Gambar 4. Plot pengelompokan kabupaten/kota di Provinsi Nusa Tenggara Timur dengan K-Means.

Berdasarkan analisis K-Means maka dapat diperoleh segmentasi kabupaten/kota yang terdiri atas 4 kelompok seperti yang ditampilkan pada *output* di bawah ini.

Clustering vector:

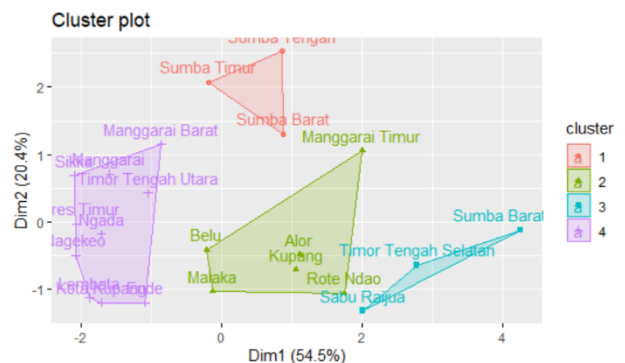
Sumba Barat	Sumba Timur	Kupang
2	2	3
Timor Tengah Selatan	Timor Tengah Utara	Belu
4	1	1
Alor	Lembata	Flores Timur
3	1	1
Sikka	Ende	Ngada
1	1	1
Manggarai	Rote Ndao	Manggarai Barat
1	3	1
Sumba Tengah	Sumba Barat Daya	Nagekeo
2	3	1
Manggarai Timur	Sabu Raijua	Malaka
3	4	1
Kota Kupang		
1		

Gambar 5. *Output* segmentasi kabupaten/kota di Provinsi Nusa Tenggara Timur dengan teknik K-Means.

Segmentasi kabupaten/kota di Provinsi Nusa Tenggara Timur menurut analisis gerombol K-Means yang terdiri atas 4 kelompok. Kelompok 1 terdiri atas 12 kabupaten/kota yaitu Kabupaten Timor Tengah Utara, Belu, Lembata, Flores Timur, Sikka, Ende, Ngada, Manggarai, Manggarai Barat, Malaka,

Nagekeo, dan Kota Kupang. Kelompok 2 hanya terdiri atas 3 kabupaten/kota yaitu Sumba Barat, Sumba Timur, dan Sumba Tengah. Kelompok 3 terdiri atas 5 kabupaten/kota yaitu Alor, Kupang, Rote Ndao, Sumba Barat Daya, dan Manggarai Timur. Kelompok 4 terdiri atas 2 kabupaten/kota yaitu Timor Tengah Selatan dan Sabu Raijua.

Selanjutnya digunakan teknik alternatif dari analisis gerombol K-Means yaitu analisis K-Medoids pada data kabupaten/kota berdasarkan indikator kesehatan dengan banyak kelompok optimal yaitu $k=4$. Plot kelompok-kelompok yang terbentuk dari hasil analisis gerombol K-Medoids ditampilkan pada Gambar 6.



Gambar 6. Plot pengelompokan kabupaten/kota di Provinsi Nusa Tenggara Timur dengan teknik K-Medoids.

Menurut analisis gerombol K-Medoids dengan banyak gerombol optimal sebanyak 4, diperoleh segmentasi kabupaten/kota di Provinsi Nusa Tenggara Timur seperti yang ditampilkan *output software* R pada Gambar 7.

Clustering vector:

Sumba Barat	Sumba Timur	Kupang
1	1	2
Timor Tengah Selatan	Timor Tengah Utara	Belu
3	4	2
Alor	Lembata	Flores Timur
2	4	4
Sikka	Ende	Ngada
4	4	4
Manggarai	Rote Ndao	Manggarai Barat
4	2	4
Sumba Tengah	Sumba Barat Daya	Nagekeo
1	3	4
Manggarai Timur	Sabu Raijua	Malaka
2	3	2
Kota Kupang		
4		

Gambar 7. *Output* segmentasi kabupaten/kota di Provinsi Nusa Tenggara Timur dengan teknik K-Medoids.

Analisis gerombol dengan teknik K-Medoids membagi data kabupaten/kota menjadi 4 gerombol berdasarkan indikator kesehatan. Pada kelompok 1 terdapat 3 kabupaten/kota yaitu Sumba Timur, Sumba Tengah, dan Sumba Barat. Kelompok 2 memiliki 5 kabupaten/kota yaitu Kupang, Belu, Alor, Manggarai Timur, dan Malaka sedangkan Kelompok 3 terdiri atas 3 kabupaten/kota yaitu Timor Tengah Selatan, Sumba Barat Daya, dan Sabu Raijua. Kabupaten/kota yang masuk pada Kelompok 4 adalah Kabupaten Lembata, Flores Timur, Sikka, Ende, Ngada, Manggarai, Manggarai Barat, Nagekeo, dan Kota Kupang.

Perbandingan Davies-Bouldin Index (DBI)

Setelah dilakukan analisis gerombol baik menggunakan K-Means maupun K-Medoids, selanjutnya akan dipilih teknik penggerombolan yang lebih baik dalam mengklasifikasikan kabupaten/kota di wilayah Provinsi Nusa Tenggara Timur berdasarkan indikator kesehatan.

Davies-Bouldin Index (DBI) digunakan sebagai nilai perbandingan dari kedua teknik penggerombolan. Semakin kecil nilai DBI maka semakin baik klasifikasi yang dilakukan.

Hasil penghitungan DBI pada teknik K-Means dan K-Medoids menurut *output* ditampilkan pada Tabel 2 berikut.

Tabel 2. Hasil penghitungan Davies-Bouldin Index bagi K-Means dan K-Medoids.

Teknik Gerombol	DBI
K-Means	1.15382
K-Medoids	1.17741

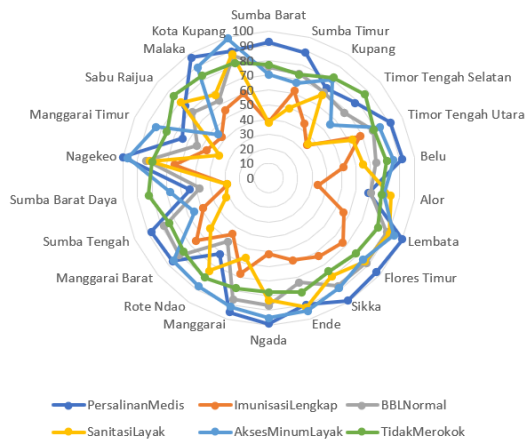
Teknik analisis gerombol menggunakan K-Means memiliki angka DBI yang lebih kecil dibandingkan dengan teknik K-Medoids, sehingga metode yang lebih baik digunakan untuk mengelompokkan kabupaten/kota di wilayah Provinsi Nusa Tenggara Timur berdasarkan indikator kesehatan adalah K-Means. Maka diperoleh 4 kelompok kabupaten/kota yang memiliki status kesehatan berbeda berdasarkan teknik K-Means yaitu:

Kelompok 1: Timor Tengah Utara, Belu, Lembata, Flores Timur, Sikka, Ende, Ngada, Manggarai, Manggarai Barat, Malaka, Nagekeo, dan Kota Kupang.

Kelompok 2: Sumba Barat, Sumba Timur, dan Sumba Tengah.

Kelompok 3: Alor, Kupang, Rote Ndao, Sumba Barat Daya, dan Manggarai Timur.

Kelompok 4: Timor Tengah Selatan dan Sabu Raijua.



Gambar 8. Plot radar sebaran indikator masing-masing kabupaten/kota.

Dari Gambar 8 dapat diperoleh karakteristik masing-masing kelompok kabupaten/kota yang terbentuk berdasarkan indikator kesehatan. Kelompok 1 merupakan kabupaten/kota di wilayah Provinsi Nusa Tenggara Timur yang memiliki nilai indikator-indikator kesehatan pada penelitian ini relatif lebih tinggi dibandingkan dengan kelompok lainnya. Kelompok 2 merupakan kabupaten/kota di Provinsi Nusa Tenggara Timur dengan status kesehatan yang cukup baik namun memiliki persentase indikator sanitasi layak yang rendah. Kelompok 3 merupakan kabupaten/kota di Provinsi Nusa Tenggara Timur dengan status kesehatan yang cukup baik namun memiliki persentase imunisasi lengkap yang rendah. Kelompok 4 memiliki karakteristik dengan nilai indikator persentase imunisasi lengkap dan persentase rumah tangga dengan sanitasi layak yang rendah dibandingkan kelompok lainnya.

4. Simpulan dan Saran

Pada penelitian ini digunakan dua teknik analisis gerombol untuk mengelompokkan wilayah kabupaten/kota di Provinsi Nusa Tenggara Timur berdasarkan indikator-indikator kesehatan yaitu K-Means dan K-Medoids dengan banyak kelompok yang optimal yaitu 4 kelompok. Berdasarkan nilai Davies-Bouldin Index (DBI) dari kedua teknik analisis gerombol, diketahui bahwa teknik K-Means lebih baik digunakan untuk mengelompokkan kabupaten/kota di Provinsi Nusa Tenggara Timur menurut status kesehatan tahun 2019.

Melalui penelitian ini ditemukan bahwa kualitas aspek kesehatan dari 22 kabupaten/kota di Provinsi Nusa Tenggara Timur belum merata. Terbentuk kelompok yang terdiri dari kabupaten/kota dengan indikator kesehatan baik, kelompok dengan kabupaten/kota yang perlu meningkatkan kebersihan sanitasi rumah tangga masyarakatnya, kelompok kabupaten/kota yang harus menggalakan imunisasi lengkap bagi anak usia 0-59 bulan dan kelompok kabupaten/kota yang perlu memperbaiki kualitas sanitasi rumah tangga beserta imunisasi anak lengkap.

Kelompok pertama berdasarkan hasil analisis memiliki anggota kabupaten/kota dengan nilai indikator kesehatan yang paling baik, kelompok kedua terdiri atas kabupaten/kota yang memiliki indikator kesehatan cukup baik namun kondisi sanitasi di rumah tangga

perlu diperbaiki. Kelompok ketiga mencakup kabupaten/kota dengan kondisi status kesehatan yang cukup baik tetapi perlu ada peningkatan persentase imunisasi lengkap bagi anak usia 0-59 bulan, dan kelompok keempat anggotanya adalah kabupaten/kota yang persentase sanitasi layak dan pemberian imunisasi lengkap anak usia 0-59 bulan yang rendah.

Penelitian ini dapat dimanfaatkan bagi pengambilan kebijakan guna tercapainya kesejahteraan masyarakat terutama pembangunan kesehatan di Provinsi Nusa Tenggara Timur. Pengelompokan yang dihasilkan dari penelitian ini dapat menjadi pertimbangan untuk meningkatkan nilai indikator-indikator kesehatan sesuai dengan karakteristik masing-masing kelompok kabupaten/kota sehingga dapat memperbaiki status kesehatan masyarakat di wilayah Nusa Tenggara Timur.

Pemerintah kabupaten/kota maupun provinsi diharapkan dapat terus mendukung dan mempertahankan program-program kesehatan pada kabupaten/kota yang berada di kelompok pertama karena menurut hasil penelitian, Kabupaten Timor Tengah Utara, Belu, Lembata, Flores Timur, Sikka, Ende, Ngada, Manggarai, Manggarai Barat, Malaka, Nagekeo, dan Kota Kupang telah memiliki indikator kesehatan yang baik. Untuk kabupaten/kota yang berada pada kelompok kedua yaitu Kabupaten Sumba Barat, Sumba Timur, dan Sumba Tengah, pemerintah diharapkan dapat

membuat atau memperbaiki kebijakan kesehatan bagi sanitasi di rumah tangga karena menurut hasil penelitian, pada kabupaten/kota di kelompok kedua ini persentase sanitasi rumah tangga yang layak masih rendah.

Pada kabupaten/kota di kelompok ketiga yaitu Kabupaten Alor, Kupang, Rote Ndao, Sumba Barat Daya, dan Manggarai Timur, pemerintah perlu memperhatikan program-program kesehatan terkait imunisasi lengkap pada anak usia 0-59 bulan karena berdasarkan hasil penelitian ini, cakupan persentase imunisasi lengkap pada anak usia 0-59 bulan di seluruh kabupaten/kota tersebut masih rendah. Sedangkan bagi wilayah yang berada di kelompok keempat yaitu Kabupaten Timor Tengah Selatan dan Sabu Raijua, pemerintah diharapkan dapat menaruh perhatian lebih pada indikator sanitasi rumah tangga dan imunisasi anak usia 0-59 bulan karena menurut hasil penelitian, persentase rumah tangga pada kabupaten Timor Tengah Selatan dan Sabu Raijua dengan sanitasi yang layak masih rendah serta persentase imunisasi lengkap yang diberikan pada anak usia 0-59 bulan juga perlu ditingkatkan.

Daftar Pustaka

- Arora, P., Deepali, D., & Varshney, S. (2016). Analysis of K-Means and K-Medoids Algorithm For Big Data. *Procedia - Procedia Computer Science*, **78**, 507–512. <https://doi.org/10.1016/j.procs.201>

- 6.02.095
 BPS. (2019). *Statistik Kesejahteraan Rakyat Provinsi Nusa Tenggara Timur 2019*. BPS Provinsi Nusa Tenggara Timur
- Debataraja, Ewaldus. (2019). PENERAPAN METODE K-MEDOIDS PADA PENGELOMPOKAN DAERAH PENGHASIL KELAPA SAWIT DENGAN VALIDASI INDEKS SILHOUETTE. *Bimaster : Buletin Ilmiah Matematika, Statistika dan Terapannya*. 8. 10.26418/bbimst.v8i4.36362.
- Dewa, F. A., & Jatipaningrum, M. T. (2019). SEGMENTASI E-COMMERCE DENGAN CLUSTER K-MEANS DAN FUZZY C-MEANS: Studi Kasus: Media Sosial di Indonesia yang diunduh di Play Store. *Jurnal Statistika Industri Dan ...*, 4(1), 53–67.
- EMC Education Services. (2015). *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Indianapolis, Indiana: John Wiley & Sons, Inc.
- Hardiani, T., Sulistyio, S., & Hartanto, R. (2015). Segmentasi Nasabah Tabungan Menggunakan Model RFM (Recency, Frequency, Monetary) dan K-Means Pada Lembaga Keuangan Mikro. *Seminar Nasional Teknologi Informasi Dan Komunikasi Terapan*, 2015.
- Izenman, A. (2008). *Modern Multivariate Statistical Techniques*. New York (US): Springer.
- Jumadi Dehotman Sitompul, B., Salim Sitompul, O., & Sihombing, P. (2019). Enhancement Clustering Evaluation Result of Davies-Bouldin Index with Determining Initial Centroid of K-Means Algorithm. *Journal of Physics: Conference Series*, 1235(1). <https://doi.org/10.1088/1742-6596/1235/1/012015>
- Kantardzic, M. (2011). *Data Mining: Concepts, Models, Methods, and Algorithms*. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Kementerian Kesehatan RI. (2013). *Riskesdas Dalam Angka Provinsi Nusa Tenggara Timur 2013*. Jakarta: Lembaga Penerbit Badan Litbangkes
- Kementerian Kesehatan RI. (2019). *Indeks Pembangunan Kesehatan Masyarakat Tahun 2018*. Jakarta: Lembaga Penerbit Badan Litbangkes.
- Kumar, P., & Sirohi, D. (2010). Comparative analysis of FCM and HCM algorithm on Iris data set. *International Journal of Computer Applications*, 5(2), 33–37. <https://doi.org/10.5120/888-1261>
- Liu, Z., Ren, C., & Cai, W. (2020). Overview of Clustering Analysis Algorithms in Unknown Protocol Recognition. *MATEC Web of Conferences CSCNS2019*, 309(03008). <https://doi.org/doi.org/10.1051/ma-teccconf/202030903008>
- Mayasari, T. R. (2019). Clustering Akses Air Bersih Dan Sanitasi Layak (Clustering of Clean Water access and Worth Sanitation in District / City Lampung Province). *Fungsional Statistisi Pertama BPS Kabupaten Pesawaran*, 563–572.
- Munawar, M., Aulia, R., Ferdhiana, R., Marzuki, M. & Iqbal, M. (2020). Identifikasi Faktor-Faktor Indeks Pembangunan Kesehatan Masyarakat Provinsi Aceh dengan Pendekatan Analisis Jalur. 20. 91-100.

- Octavanny, M.A.D. & Budiantara, I. & Ratnasari, V. (2017). Pemodelan Faktor-Faktor yang Memengaruhi Indeks Pembangunan Kesehatan Masyarakat Provinsi Jawa Timur Menggunakan Pendekatan Regresi Semiparametrik Spline. *Jurnal Sains dan Seni ITS*. 6. 10.12962/j23373520.v6i1.22491.
- Putri, M. M. & Fithriasari, K. (2015). Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan Masyarakat Menggunakan Metode Kohonen SOM dan *K-Means*. *Jurnal Sains dan Seni ITS Vol. 4 No. 1*.
- Salazar, D. A., Velez, J. I., & Salazar, J. C. (2012). Comparison between SVM and Logistic Regression: Which One is Better to Discriminate? *Revista Colombiana de Estadística Número Especial En Bioestadística*, 35(2), 223–237. Retrieved from <http://emis.impa.br/EMIS/journals/RCE/V35/v35n2a03.pdf>
- Soni, K. G., & Patel, A. (2017). Comparative Analysis of K-means and K-medoids Algorithm on IRIS Data. *International Journal of Computational Intelligence Research*, 13(5), 899–906. Retrieved from https://www.ripublication.com/ijci_r17/ijcirv13n5_21.pdf
- Thinsungnoen, T., Kaoungku, N., Durongdumronchai, P., Kerdprasop, K., & Kerdprasop, N. (2015). The Clustering Validity with Silhouette and Sum of Squared Errors, 44–51. <https://doi.org/10.12792/iciae2015.012>
- Thrun, M. C. (2017). *Projection-Based Clustering through Self-Organization and Swarm Intelligence*. Marburg, Germany: Springer Viewwg. <https://doi.org/10.1007/978-3-658-20540-9>
- Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 411–423. <https://doi.org/10.1111/1467-9868.00293>
- Wulandari, S. (2020). Clustering Microarray Adenoma Menggunakan Spectral Clustering dengan Algoritma Partitioning Around Medoid (PAM). *Sinasis I*, 1(1), 345–351.
- Yuan, C., & Yang, H. (2019). Research on K-Value Selection Method of K-Means Clustering Algorithm. *J*, 2(2), 226–235. <https://doi.org/10.3390/j2020016>