

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

Pengelompokkan Kabupaten/Kota di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan K-Medoids Clustering

Sitasi : Wicaksono & Yolanda, 2021, JSTAR 1(1), 79-90.

Kronologi naskah.

Submit : 22 Mei 2021

Revisi : 31 Mei 2021

Diterima : 28 Juni 2021



Penyedia Data Statistik Berkualitas untuk
Indonesia Maju

REFORMASI BIROKRASI



Pengelompokkan Kabupaten/Kota di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan *K-Medoids Clustering*

Agung Satrio Wicaksono¹, Anne Mudya Yolanda²

¹Program Studi Administrasi Publik, Fakultas Ilmu Sosial dan Ilmu Politik,

Universitas Sultan Ageng Tirtayasa, Indonesia

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

*korespondensi author: agungsatriow@untirta.ac.id¹, annemudyayolanda@lecturer.unri.ac.id²

Abstract

The Human Development Index (HDI) is an indicator to measure success in improving the quality of human life. The utilization of HDI indicators can be used to classify observations into clusters based on several aspects, such as health, education, and economy. The results of the clusters can be used as a reference for evaluating government policies. This study uses HDI indicator data in the East Nusa Tenggara (NTT) Province, which consists: Life Expectancy, Average Years of Schooling, Expected Years of Schooling, and Per Capita Expenditure. The outliers were detected in each variable. Therefore the k-medoids clustering method was used in this study because of their robustness on outliers. The clustering results show that there are 4 clusters formed, with each cluster describing a unique character. Cluster 1 describes the condition of the districts in general. Cluster 2 describes the condition of the districts that are one step ahead of others, which are expected to transform into cities in the future. Cluster 3 describes the condition of the district that needs more attention from the government in all sectors because all the HDI indicators are on the lowest value comparing to the other clusters. On the other hand, cluster 4 describes urban areas which have more developed than other areas, in terms of health, education, and economy. The clustering results can be used as a suggestion for evaluating government policies that have been implemented, in order to create structured equitable development based on data and regional characteristics as well.

Keywords: *clustering, human development index (HDI) indicator, k-medoids.*

1. Pendahuluan

Meningkatnya pemahaman tentang pentingnya dimensi non-ekonomi pada ukuran kesejahteraan masyarakat ditandai dengan keterlibatan dimensi sosial, politik dan budaya dalam pengukuran kesejahteraan masyarakat (Badrudin, 2012). Salah satu ukuran

yang mengakomodir dimensi-dimensi non-ekonomi dalam pengukuran kesejahteraan masyarakat tersebut adalah konsep tentang indeks pembangunan manusia, yang diintroduksi oleh *United Nations Development Program*.

Indeks pembangunan manusia

(*Human Development Index*) merupakan sebuah indikator penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia. Indeks pembangunan manusia dibentuk oleh tiga dimensi dasar kehidupan manusia, yaitu kesehatan, pendidikan serta ekonomi (Kpolovie dkk., 2017). Indeks pembangunan manusia dapat menentukan peringkat atau level pembangunan suatu wilayah. Ketiga dimensi dasar yang membentuk nilai indeks pembangunan manusia diukur oleh masing-masing indikator yang mewakilinya.

Indikator yang mengukur dimensi kesehatan yaitu angka harapan hidup (*life expectancy*). Secara umum, angka harapan hidup merupakan sebuah indikator untuk mengevaluasi kinerja pemerintah dalam peningkatan kesejahteraan hidup masyarakat (Sinaga dkk., 2020). Menurut definisi, angka harapan hidup merupakan rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang sejak lahir. Idealnya angka harapan hidup dihitung berdasarkan angka kematian menurut umur (*Age Specific Death Rate/ASDR*) yang datanya diperoleh dari catatan registrasi kematian secara bertahun-tahun sehingga dimungkinkan dibuat tabel kematian. Sumber data lain untuk mendapatkan variabel penyusun indikator ini adalah sensus penduduk dan survei penduduk antar sensus (SUPAS). Angka harapan hidup mewakili dimensi kesehatan sebagai salah satu indikator perhitungan Indeks Pembangunan Manusia.

Dimensi kedua yaitu dimensi

pendidikan, dimana dimensi ini terdiri dari dua indikator, yaitu rata-rata lama sekolah (*mean years of schooling*) dan harapan lama sekolah (*expected years of schooling*). Rata-rata lama sekolah didefinisikan sebagai jumlah tahun yang digunakan oleh penduduk dalam menjalani pendidikan formal, diasumsikan bahwa dalam kondisi normal rata-rata lama sekolah suatu wilayah tidak akan turun (Chamidah dkk., 2020). Cakupan penduduk yang dihitung rata-rata lama sekolahnya adalah penduduk berusia 25 tahun ke atas dengan asumsi bahwa pada umur 25 tahun proses pendidikan sudah berakhir. Penghitungan ini mengikuti standar internasional yang digunakan oleh *United Nations Development Program*. Sedangkan, harapan lama sekolah didefinisikan sebagai lamanya sekolah (dalam tahun) yang diharapkan akan dirasakan oleh anak pada umur tertentu di masa mendatang. Dalam hal ini diasumsikan bahwa peluang anak tersebut akan tetap bersekolah pada umur-umur berikutnya sama dengan peluang penduduk yang bersekolah per jumlah penduduk untuk umur yang sama saat ini (Masuda & Yudhistira, 2020). Angka harapan lama sekolah dihitung untuk penduduk berusia 7 tahun keatas karena mengikuti kebijakan pemerintah yaitu usia minimum untuk melaksanakan program wajib belajar. Angka ini dapat digunakan untuk mengetahui kondisi pembangunan sistem pendidikan di berbagai jenjang yang ditunjukkan dalam bentuk lamanya pendidikan (dalam tahun) yang diharapkan dapat dicapai oleh setiap anak.

Dimensi ketiga yaitu dimensi ekonomi, dengan indikatornya yaitu pengeluaran per kapita yang disesuaikan, yaitu biaya yang dikeluarkan untuk konsumsi semua anggota rumah tangga selama sebulan dibagi dengan banyaknya anggota rumah tangga. Data pengeluaran dapat mengungkap tentang pola konsumsi rumah tangga secara umum menggunakan indikator proporsi pengeluaran untuk makanan dan non makanan. Komposisi pengeluaran rumah tangga dapat dijadikan ukuran untuk menilai tingkat kesejahteraan ekonomi penduduk, makin rendah persentase pengeluaran untuk makanan terhadap total pengeluaran makin membaik tingkat kesejahteraan.

Dalam pemanfaatannya, indikator indeks pembangunan manusia dapat digunakan untuk mengelompokkan wilayah berdasarkan aspek kesehatan, pendidikan dan ekonomi. Pengelompokan tersebut perlu dilakukan sebagai bahan perencanaan dan evaluasi sasaran program pemerintah (Kiha dkk., 2021). Tujuan lain dari pengelompokan wilayah tersebut yaitu untuk meningkatkan angka pembangunan manusia berdasarkan indikator pembentuk indeks pembangunan manusia. Salah satu teknik analisis statistika yang digunakan untuk mengelompokkan data berdasarkan karakteristik peubahnya secara alami yaitu analisis gerombol (*cluster analysis*).

Analisis gerombol (*cluster analysis*) merupakan suatu metode peubah ganda

yang bertujuan untuk mengelompokkan n objek ke dalam k kelompok berdasarkan karakter-karakternya, di mana $k \leq n$ (Johnson dkk., 2002). Objek-objek yang berada di dalam kelompok yang sama akan memiliki tingkat kesamaan yang tinggi dibandingkan objek-objek antar kelompok yang berbeda. Terdapat dua metode dasar dalam analisis gerombol, yaitu metode hirarki yang menyangkut dengan pembentukan struktur hirarki (*dendogram*) dan metode non-hirarki yang digunakan apabila banyak gerombol k sudah ditentukan terlebih dahulu. Penentuan k dapat ditentukan secara subjektif peneliti maupun objektif dengan menggunakan metode *silhouette coefficient*, metode *elbow*, maupun metode lainnya.

Algoritma *k-medoids clustering* yang juga sering disebut *partitioning around medoids* (PAM) adalah sebuah metode penggerombolan non-hirarki yang merupakan variasi dari metode *k-means*. *K-medoids* dapat mengatasi kelemahan *k-means* yang cenderung sensitif terhadap pencilan (*outliers*) yang mungkin menyimpang dari distribusi data (Schubert & Rousseeuw, 2019). Berbeda dengan *k-means* yang menggunakan rata-rata sebagai titik pusat gerombol, *k-medoids* menggunakan medoid sebagai titik pusat gerombol. Medoid merupakan objek yang letaknya terpusat di dalam suatu gerombol, sehingga lebih kekar terhadap pencilan. Gerombol dibangun dengan menghitung kedekatan jarak antara medoid dengan objek non-

medoid (Sangga, 2018).

Pada penelitian ini dibahas tentang pemanfaatan algoritma *k-medoids clustering* untuk mengelompokkan kabupaten/kota di Provinsi Nusa Tenggara Timur berdasarkan karakteristik yang diukur dari indikator indeks pembangunan manusia, yaitu angka harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran per kapita yang disesuaikan. Penentuan banyaknya gerombol k ditentukan secara objektif menggunakan metode *silhouette coefficient*. Secara umum, hasil yang diperoleh dapat dimanfaatkan sebagai acuan bahan perencanaan dan evaluasi sasaran program pemerintah, khususnya yang berkaitan dengan kesejahteraan masyarakat.

2. Metodologi

Sumber data

Penelitian ini menggunakan data yang bersumber dari Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur (NTT). Data yang digunakan adalah data indikator indeks pembangunan manusia tahun 2020, dengan peubah yang terdiri dari angka harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pengeluaran per kapita yang disesuaikan. Hasil pengelompokkan diharapkan dapat menjadi bahan perencanaan dan evaluasi sasaran program pemerintah.

Metode Penelitian

Langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut:

- 1) Eksplorasi data, yang bertujuan untuk melihat karakteristik dan sebaran dari data.
- 2) Penentuan banyaknya gerombol (k) secara objektif menggunakan metode *silhouette coefficient (SC)*. Metode ini merupakan gabungan dari metode separasi dan kohesi (Dinh dkk., 2019). Berikut merupakan rumus umum untuk menentukan SC:

$$SC = \max_k SI(k) \quad (1)$$

di mana k merupakan banyaknya gerombol, dan SI merupakan nilai *silhouette index*, yang didapatkan melalui perhitungan berikut:

$$SI = \frac{1}{k} \sum_{j=1}^k SI_j \quad (2)$$

di mana SI_j merupakan rata-rata *silhouette index* gerombol- j , yang didapatkan melalui perhitungan berikut:

$$SI_j = \frac{1}{m_j} \sum_{i=1}^{m_j} SI_i^j \quad (3)$$

di mana m_j merupakan jumlah data dalam gerombol ke- j , i merupakan index data ($i = 1, 2, \dots, m_j$), dan SI_i^j merupakan *silhouette index* data ke- i dalam gerombol- j , yang didapatkan melalui perhitungan berikut:

$$SI_i^j = \frac{b_i^j - a_i^j}{\max\{a_i^j, b_i^j\}} \quad (4)$$

di mana b_i^j merupakan rata-rata jarak data ke- i terhadap semua data yang tidak dalam satu gerombol dengan data ke- i , dan a_i^j merupakan

rata-rata jarak data ke- i terhadap semua data dalam satu gerombol, yang didapatkan melalui perhitungan berikut:

$$b_i^j = \min_{\substack{n=1, \dots, k \\ n \neq j}} \left\{ \frac{1}{m_n} \sum_{\substack{r=1 \\ i \neq r}}^{m_n} d(x_i^j, x_r^n) \right\} \quad (5)$$

$$a_i^j = \frac{1}{m_j - 1} \sum_{\substack{r=1 \\ i \neq r}}^{m_j} d(x_i^j, x_r^j) \quad (6)$$

di mana j adalah gerombol, m_j adalah banyaknya data dalam gerombol ke- j , $d(x_i^j, x_r^j)$ adalah jarak data ke- i dengan data ke- r dalam satu gerombol- j , m_n adalah banyaknya data dalam gerombol ke- n , dan $d(x_i^j, x_r^n)$ adalah jarak data ke- i dengan data ke- r dengan satu gerombol- n .

3) Lakukan *k-medoids clustering* dengan k yang didapatkan dari hasil metode SC. Algoritma *k-medoids* adalah sebagai berikut:

- (i) Secara acak, pilih k objek sebagai medoid awal, di mana setiap objek tersebut dipandang sebagai sebuah gerombol yang berbeda.
- (ii) Hitung jarak setiap objek non-medoid ke medoid awal tersebut, kemudian tandai jarak ke medoid terdekat, dan hitung total jarak terdekatnya.
- (iii) Secara acak, pilih k objek non-medoid sebagai kandidat medoid baru.
- (iv) Hitung jarak setiap objek

dengan kandidat medoid baru, kemudian tandai jarak ke medoid terdekat, dan hitung total jarak terdekatnya

- (v) Hitung total simpangan (S) dengan mengurangi total jarak terdekat pada medoid awal (poin ii) dengan total jarak terdekat pada kandidat medoid baru (poin iv).
- (vi) Jika hasil $S < 0$, jadikan kandidat medoid baru sebagai medoid awal, kemudian lakukan iterasi dengan mengulangi langkah (iii) sampai (v). Jika hasil $S \geq 0$, proses selesai dan dapat dihentikan.

Jarak dari setiap objek non-medoid ke medoid dihitung menggunakan ukuran jarak euclid (*euclidean distance*), yang dirumuskan sebagai berikut:

$$d_{euc}(x_{ij}, c_{kj}) = \sqrt{\sum_{j=1}^p \sum_{i=1}^n (x_{ij} - c_{kj})^2} \quad (7)$$

dimana $d_{euc}(x_{ij}, c_{kj})$ merupakan jarak euclid antara objek- i pada peubah ke- j ke pusat gerombol- k pada peubah ke- j , x_{ij} merupakan objek- i pada peubah ke- j , c_{kj} merupakan pusat gerombol- k pada peubah ke- j , p merupakan banyaknya peubah yang diamati, dan n merupakan banyaknya objek.

- 4) Interpretasi hasil penggerombolan, dilakukan untuk mengetahui karakteristik dari setiap gerombol yang terbentuk. Interpretasi hasil

penggerombolan dapat dilihat dari rata-rata peubah pada setiap gerombol yang terbentuk.

3. Hasil dan Pembahasan

Pengolahan data dilakukan dengan bantuan *software* R. Berikut merupakan hasil yang didapatkan.

Eksplorasi Data

Langkah pertama dalam eksplorasi data yaitu memeriksa *missing value*. *Missing value* atau data hilang dapat mengakibatkan hasil pengolahan data menjadi bias. Hasil pengecekan dilakukan dengan bantuan *software* R. Hasil *output* pengecekan data hilang ditampilkan dalam Gambar 1.

```

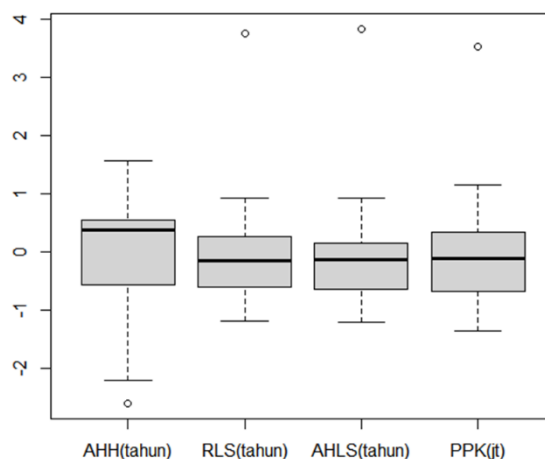
AHH(tahun)   RLS(tahun)
Mode :logical Mode :logical
FALSE:22     FALSE:22
AHL5(tahun)  PPK(jt)
Mode :logical Mode :logical
FALSE:22     FALSE:22

```

Gambar 1. Hasil cek *missing value*.

Dari Gambar 1, terlihat bahwa nilai FALSE:22 menunjukkan bahwa dari 22 amatan, tidak terdapat kasus *missing value*. Artinya data sudah bebas dari data hilang.

Setelah memeriksa data hilang, dilanjutkan dengan pemeriksaan outlier atau pencilan. Hasil *boxplot* pada Gambar 2 menjelaskan bahwa terdapat pencilan pada data. Pencilan merupakan amatan yang menyimpang cukup jauh dari amatan yang lainnya, sehingga menyebabkan dugaan bahwa amatan yang menyimpang tersebut berasal dari sebaran data yang berbeda (Zulkarnain dkk. 2020).



Gambar 2. *Boxplot* hasil standarisasi data.

Pada Gambar 2, terlihat bahwa terdapat sebuah amatan pencilan pada setiap peubahnya. Secara umum, untuk mengatasi pencilan, amatan dapat dibuang atau dipertahankan, tergantung kepada pengambilan keputusan yang dilakukan oleh peneliti dengan pertimbangan yang berdasarkan objek penelitian. Dalam penelitian ini, pencilan tidak perlu dihilangkan karena metode penggerombolan k-medoids yang digunakan dapat kekar terhadap pencilan. Selain itu, satu baris data mempresentasikan keadaan satu kabupaten/kota di Provinsi NTT.

Tabel 1. Statistik deskriptif.

Statistik	AHH	RLS	HLS	PPK
Min	60.64	6.25	11.99	5.26
Q1	65.01	6.89	12.50	6.41
Median	67.00	7.36	12.94	7.31
Rataan	66.15	7.52	13.04	7.50
Q3	67.34	7.81	13.18	7.99
Maks	69.55	11.58	16.40	13.34

Keterangan:

AHH : Angka Harapan Hidup (Tahun)
 RLS : Rata-rata Lama Sekolah (Tahun)
 HLS : Harapan Lama Sekolah (Tahun)
 PPK : Pengeluaran Per Kapita yang disesuaikan (Juta Rupiah)

Hasil perhitungan statistik deskriptif dari masing-masing peubah disajikan dalam Tabel 1. Hasil perhitungan statistik deskriptif menunjukkan bahwa rata-rata angka harapan hidup penduduk NTT yaitu 66.15 tahun, dengan nilai terendah yaitu 60.64 yang berada di wilayah Kabupaten Sabu Raijua. Sedangkan nilai tertinggi yaitu 69.55 yang berada di wilayah Kota Kupang. Terdapat perbedaan selisih 8.91 tahun antara wilayah Kabupaten Sabu Raijua dengan Kota Kupang.

Selanjutnya, didapatkan rata-rata nilai rata-rata lama sekolah yaitu 7.52 tahun. Artinya, penduduk berusia 25 tahun ke atas rata-rata telah menghabiskan waktu 7.53 tahun dalam menjalani pendidikan formal. Lamanya waktu tersebut menunjukkan bahwa rata-rata penduduk NTT yang berusia 25 tahun keatas telah menjalani pendidikan sampai dengan tingkat kedua di Sekolah Menengah Pertama (SMP). Rata-rata lama sekolah tersebut faktanya masih di bawah standar lamanya wajib belajar yang ditetapkan oleh Pemerintah. Jika ditelusuri ke belakang, pada tahun 1994 Pemerintah kala itu telah menerapkan program wajib belajar 9 tahun. Artinya, penduduk yang pada tahun 2020 telah berumur 25 tahun ke atas seharusnya telah menjalani pendidikan formal selama minimal 9 tahun.

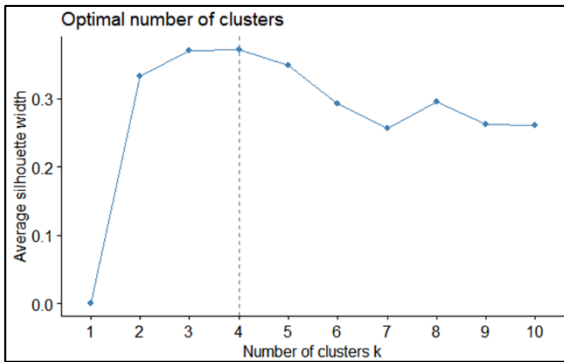
Hasil rata-rata nilai harapan lama sekolah yang didapatkan yaitu 13.04. Artinya, pada saat usia 25 tahun, penduduk yang saat ini berusia 7 tahun keatas diharapkan telah menjalani pendidikan formal selama minimal

13.04 tahun. Hal tersebut selaras dengan kebijakan Pemerintah yang mulai pada tahun 2015 merevisi program wajib belajar 9 tahun menjadi 12 tahun. Program tersebut tentunya juga perlu ditunjang dengan kemampuan guru dan fasilitas sekolah yang memadai, termasuk di Provinsi NTT.

Dari dimensi ekonomi, nilai rata-rata pengeluaran per kapita yang disesuaikan yaitu 7.50 juta rupiah. Artinya, rata-rata biaya yang digunakan untuk konsumsi setiap anggota keluarga penduduk NTT selama tahun 2020 adalah sebesar 7.50 juta rupiah. Daerah dengan pengeluaran per kapita terendah sebesar 5.26 juta rupiah yaitu Kabupaten Sabu Raijua, sedangkan yang tertinggi yaitu daerah Kota Kupang, dengan pengeluaran per kapita sebesar 13.34 juta rupiah. Terdapat perbedaan selisih sebesar 8.08 juta rupiah antara wilayah Kabupaten Sabu Raijua dengan Kota Kupang.

Penentuan Banyaknya Gerombol

Langkah selanjutnya yaitu menentukan banyaknya gerombol (k) secara objektif dengan menggunakan metode *silhouette coefficient*. Hasil perhitungan metode *silhouette coefficient* didapatkan dengan bantuan *software R* yang ditampilkan dalam bentuk grafik pada Gambar 3. Pendekatan nilai rata-rata digunakan untuk menduga hasil kualitas gerombol, dimana nilai rata-rata yang semakin tinggi berarti akan semakin baik kualitas gerombol yang dihasilkan.



Gambar 3. Grafik rata-rata untuk setiap k dari metode *Silhouette Coefficient*.

Berdasarkan hasil yang ditampilkan pada Gambar 3, didapatkan rekomendasi penentuan jumlah gerombol optimum yang dapat terbentuk yaitu $k=4$, yang merupakan nilai dengan rata-rata *silhouette coefficient* tertinggi dari yang lainnya. Penentuan banyaknya gerombol (k) ini diharapkan mampu mendapatkan hasil penggerombolan yang paling optimum.

Melakukan *k-Medoids Clustering*

Setelah menetapkan banyaknya gerombol (k), langkah berikutnya adalah melakukan penggerombolan dengan metode *k-medoids clustering*. Dengan menggunakan bantuan *software R*, dilakukan penggerombolan untuk $k=4$. Pada Tabel 2, disajikan nilai titik pusat atau medoid dari masing-masing gerombol, yang didapatkan dari hasil penggerombolan.

Tabel 2. Titik pusat gerombol.

k	Medoid	AHH	RLS	HLS	PPK
1	15	67.38	7.30	12.28	7.49
2	9	65.20	7.71	12.91	7.63
3	20	60.64	6.65	13.14	5.27
4	22	69.55	11.58	16.40	13.34

Keterangan:

AHH : Angka Harapan Hidup (Tahun)

RLS : Rata-rata Lama Sekolah (Tahun)

HLS : Harapan Lama Sekolah (Tahun)

PPK : Pengeluaran Per Kapita yang disesuaikan (Juta Rupiah)

Berdasarkan hasil yang ditampilkan pada Tabel 2, didapatkan bahwa amatan medoid atau titik pusat gerombol ke-1 adalah amatan ke-15, yaitu wilayah Kabupaten Manggarai Barat. Sedangkan untuk gerombol ke-2, ke-3, dan ke-4, masing-masing medoid atau titik pusat gerombolnya adalah amatan ke-9, ke-20, dan amatan ke-22. Masing-masing amatan tersebut adalah wilayah Kabupaten Flores Timur, Kabupaten Sabu Raijua dan Kota Kupang.

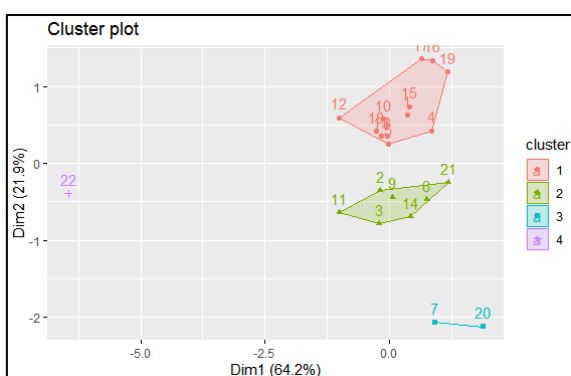
Hasil keanggotaan gerombol secara lengkap disajikan dalam Tabel 3. Diketahui terdapat 12 wilayah yang tergabung dalam gerombol ke-1, 7 wilayah dalam gerombol ke-2, 2 wilayah dalam gerombol ke-3, serta 1 wilayah dalam gerombol ke-4. Wilayah ini terbentuk secara alami berdasarkan kemiripan yang diukur dari jarak setiap titik non-medoid dengan titik medoid terdekat. Objek non-medoid yang terdekat dengan setiap medoid akan bergabung ke dalam satu kesatuan gerombol.

Tabel 3. Keanggotaan setiap gerombol

k	Wilayah
1	Kabupaten: Sumba Barat, Timor Tengah Selatan, Timor Tengah Utara, Lembata, Sikka, Ngada, Manggarai, Manggarai Barat, Sumba Tengah, Sumba Barat Daya, Nagekeo, Manggarai Timur

k	Wilayah
2	Kabupaten: Sumba Timur, Kupang, Belu, Flores Timur, Ende, Rote Ndao, Malaka
3	Kabupaten: Alor, Sabu Raijua
4	Kota: Kupang

Selanjutnya, kemiripan antar gerombol dilihat digambarkan pada hasil plot dari gerombol yang dibuat dengan bantuan *software* R. Hasil plot dari gerombol disajikan pada Gambar 4.



Gambar 4. Visualisasi hasil k-medoids.

Dari Gambar 4, didapatkan hasil visualisasi yang dapat melihat seberapa jauh perbedaan antar gerombol. Terlihat bahwa gerombol 4 merupakan gerombol yang paling berbeda dengan yang lainnya. Hal ini terjadi karena gerombol 4 memiliki jarak yang paling dengan gerombol lainnya. Sedangkan gerombol 1 dan gerombol 2 memiliki kedekatan yang paling tinggi diantara yang lainnya.

Interpretasi Hasil Penggerombolan

Karakteristik dari setiap gerombol yang terbentuk dapat dilihat pada interpretasi hasil dari setiap gerombol. Hal ini dapat dilihat dari hasil perhitungan nilai rata-rata peubah pada setiap gerombol, yang disajikan pada

Tabel 4.

Tabel 4. Rata-rata peubah pada setiap gerombol.

k	AHH	RLS	HLS	PPK	n
1	67.4	7.25	12.8	7.17	12
2	64.9	7.40	13.1	7.65	7
3	61.1	7.53	12.7	6.03	2
4	69.6	11.6	16.4	13.3	1

Keterangan:

k : Gerombol

AHH : Angka Harapan Hidup (Tahun)

RLS : Rata-rata Lama Sekolah (Tahun)

HLS : Harapan Lama Sekolah (Tahun)

PPK : Pengeluaran Per Kapita yang disesuaikan (Juta Rupiah)

n : Banyaknya wilayah dalam gerombol

Berikut merupakan interpretasi dari setiap gerombol berdasarkan hasil yang disajikan dalam Tabel 4.

- Gerombol 1 merupakan gerombol dengan cakupan wilayah terbanyak, yaitu sebanyak 12 wilayah. Gerombol ini memiliki nilai angka harapan hidup yang merupakan tertinggi kedua dibandingkan dengan gerombol lainnya sedangkan nilai rata-rata lama sekolah, harapan lama sekolah, dan pendapatan per kapita yang disesuaikan tidak lebih tinggi dari gerombol 2.
- Gerombol 2 merupakan gerombol dengan nilai rata-rata lama sekolah, harapan lama sekolah, dan pendapatan per kapita yang disesuaikan tertinggi kedua setelah

gerombol 1. Melihat fakta tersebut, diharapkan wilayah Kabupaten yang masuk ke dalam gerombol ini dapat berkembang cepat untuk bertransformasi menjadi kota di kemudian hari. Untuk mencapai hal itu, sektor kesehatan perlu lebih diperhatikan, mengingat nilai angka harapan hidup yang merepresentasikannya masih berada dibawah gerombol 1. Angka ini merupakan alat untuk mengevaluasi kinerja pemerintah yang umumnya dalam meningkatkan kesejahteraan penduduk dan khususnya dalam meningkatkan derajat kesehatan.

- c. Gerombol 3 merupakan gerombol dengan nilai angka harapan hidup, nilai rata-rata lama sekolah, harapan lama sekolah, dan pendapatan per kapita yang disesuaikan terendah jika dibandingkan dengan gerombol lainnya. Angka harapan hidup yang rendah di suatu wilayah harus diikuti dengan program pembangunan kesehatan dan program sosial lainnya termasuk kesehatan lingkungan, kecukupan gizi dan kalori, termasuk program pemberantasan kemiskinan. Angka harapan lama sekolah yang kecil menggambarkan kondisi pembangunan sistem pendidikan di berbagai jenjang. Untuk itu, wilayah yang termasuk ke dalam gerombol 1 perlu mendapat perhatian khusus pada bidang ini, mengingat pendidikan adalah hal

yang sangat penting dalam kemajuan pembangunan masa depan bagi daerah. Untuk itu, dalam mencapai pemerataan pembangunan, wilayah yang termasuk dalam gerombol ini perlu mendapat perhatian lebih, baik dari sektor kesehatan, pendidikan, hingga ekonomi.

- d. Gerombol 4 merupakan gerombol dengan nilai angka harapan hidup, rata-rata lama sekolah, harapan lama sekolah dan pendapatan per kapita yang disesuaikan yang paling tinggi daripada gerombol yang lainnya. Gerombol ini hanya terdiri dari satu amatan, yang juga merupakan satu-satunya wilayah Kota di Provinsi NTT, yaitu Kota Kupang. Hal yang wajar apabila gerombol 4 memiliki nilai tertinggi di hampir setiap indikatornya. Secara karakteristik dan demografi, wilayah perkotaan tentu akan lebih unggul dibandingkan dengan wilayah yang bukan perkotaan. Selain itu, perbedaan tersebut juga tampak dari faktor Kota Kupang yang juga menjadi ibukota Provinsi NTT.

4. Simpulan

Konsep indeks pembangunan manusia adalah salah satu ukuran yang mengakomodir dimensi-dimensi non-ekonomi dalam pengukuran kesejahteraan masyarakat. Indikator IPM terdiri dari angka harapan hidup, rata-rata lama sekolah, harapan lama sekolah, dan pendapatan per kapita yang disesuaikan. Indikator IPM dapat

menjadi indikator pengelompokan wilayah untuk meningkatkan angka pembangunan manusia. Melalui *k-Medoids Clustering* dengan peubah indikator IPM, wilayah Provinsi NTT dapat dikelompokkan ke dalam empat (4) gerombol. Gerombol 1 merupakan gerombol yang menggambarkan kondisi wilayah Kabupaten pada umumnya. Gerombol 2 merepresentasikan kumpulan wilayah dengan potensi untuk bertransformasi menjadi Kota di kemudian hari, dilihat dari tingkat rata-rata lama sekolah, harapan lama sekolah dan pengeluaran yang bernilai tertinggi kedua. Sedangkan gerombol 3 merupakan wilayah dengan angka harapan hidup, rata-rata lama sekolah, harapan lama sekolah dan pendapatan per kapita yang disesuaikan terendah dibandingkan yang lainnya, yang menandakan bahwa perlunya perhatian yang lebih dari pemerintah, baik dari sektor kesehatan, pendidikan, hingga sektor ekonomi. Gerombol 4 merepresentasikan wilayah Kota yang sudah lebih unggul dibandingkan dengan wilayah lainnya, baik dari segi angka harapan hidup, rata-rata lama sekolah, harapan lama sekolah, maupun pendapatan per kapita yang disesuaikan.

Hasil penelitian ini dapat digunakan sebagai saran untuk mengevaluasi kebijakan pemerintah yang telah diterapkan, demi terwujudnya pemerataan pembangunan yang lebih rapi dan terstruktur berdasarkan data serta karakteristik wilayah. Selain itu, penelitian berikutnya dapat

menambahkan indikator lain yang berkaitan dengan tingkat kesejahteraan. Sehingga dapat diketahui hasil pengelompokan atau penggerombolan daerah di Provinsi NTT dengan lebih menyeluruh, khususnya dari aspek lain yang belum disebutkan dalam penelitian ini. Metode statistika lain juga dapat digunakan, khususnya metode penggerombolan non-hirarki lainnya yang kekar terhadap pencilan. Hal ini dikarenakan terdapat perbedaan yang cukup signifikan antara wilayah Kabupaten dan wilayah Kota.

Daftar Pustaka

- Badrudin, R. (2012). *Pengaruh Desentralisasi Fiskal Terhadap Belanja Modal, Pertumbuhan Ekonomi, Dan Kesejahteraan Masyarakat Kabupaten/Kota Di Provinsi Jawa Tengah*. Universitas Airlangga.
- Chamidah, N., Mardianto, M. F. F., Limanta, E. E., & Hastuti, D. R. (2020). Modelling Of Poverty Percentage Based On Mean Years Of Schooling In Indonesia Using Local Linear Estimator. *The 2nd International Seminar On Science And Technology (ISSTEC 2019)*, 87–91.
- Dinh, D.-T., Fujinami, T., & Huynh, V.-N. (2019). Estimating The Optimal Number Of Clusters In Categorical Data Clustering By Silhouette Coefficient. *International Symposium On Knowledge And Systems Sciences*, 1–17.

- Johnson, R. A., Wichern, D. W., & Others. (2002). *Applied Multivariate Statistical Analysis* (Vol. 5, Issue 8). Prentice Hall Upper Saddle River, NJ.
- Kiha, E. K., Seran, S., & Lau, H. T. (2021). Pengaruh Jumlah Penduduk, Pengangguran, Dan Kemiskinan Terhadap Indeks Pembangunan Manusia (IPM) Di Kabupaten Belu. *Jurnal Ekonomi, Sosial & Humaniora*, 2(07), 60–84.
- Kpolovie, P. J., Ewansiha, S., & Esara, M. (2017). Continental Comparison Of Human Development Index (HDI). *International Journal Of Humanities Social Sciences And Education (IJHSSE)*, 4(1), 9–27.
- Masuda, K., & Yudhistira, M. H. (2020). Does Education Secularize The Islamic Population? The Effect Of Years Of Schooling On Religiosity, Voting, And Pluralism In Indonesia. *World Development*, 130, 104915.
- Sangga, V. A. P. (2018). *Perbandingan Algoritma K-Means Dan Algoritma K-Medoids Dalam Pengelompokan Komoditas Peternakan Di Provinsi Jawa Tengah Tahun 2015*.
- Schubert, E., & Rousseeuw, P. J. (2019). Faster K-Medoids Clustering: Improving The PAM, CLARA, And CLARANS Algorithms. *International Conference On Similarity Search And Applications*, 171–187.
- Sinaga, S. P., Wanto, A., & Solikhun, S. (2020). Implementasi Jaringan Syaraf Tiruan Resilient Backpropagation Dalam Memprediksi Angka Harapan Hidup Masyarakat Sumatera Utara. *Jurnal Infomedia: Teknik Informatika, Multimedia & Jaringan*, 4(2), 81–88.
- Zulkarnain, A., Rizki, S. W., & Perdana, H. (2020). Analisis Regresi Robust Estimasi-Mm Dalam Mengatasi Pencilan Pada Regresi Linear Berganda. *Bimaster*, 9(1).