

Pengelompokan Provinsi di Indonesia Berdasarkan Indikator Keluarga Sehat Menggunakan Metode Klaster Hirarki dan Non Hirarki

Ika Nur Laily Fitriana

Departemen Statistika, Fakultas Sains dan Analitika Data,
Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

*Penulis Korespondensi:
email: ikanurlaily97@gmail.com

Abstrak

Pembangunan kesehatan merupakan isu penting dan fundamental yang harus segera dilakukan oleh negara Indonesia. Tujuan penelitian ini adalah untuk mengelompokan provinsi di Indonesia berdasarkan indikator keluarga sehat yang bersumber dari Profil Kesehatan Indonesia (Kementerian Kesehatan) tahun 2019 dan Statistik Indonesia (Badan Pusat Statistik) tahun 2020. Pengelompokan dilakukan untuk mengetahui provinsi-provinsi mana yang memiliki kualitas kesehatan rendah sehingga diperlukan perbaikan akses dan fasilitas kesehatan yang lebih memadai. Pada penelitian ini dilakukan perbandingan metode klaster hirarki dan non hirarki. Metode yang terbaik antara klaster hirarki dan non hirarki dipakai dalam pengelompokan provinsi berdasarkan indikator keluarga sehat. Pemilihan metode terbaik berdasarkan nilai icdrate dan pseudo-f statistics. Metode terbaik antara klaster hirarki dan non hirarki yaitu metode k-means dengan 4 klaster dihasilkan nilai Pseudo-F statistics sebesar 13,57 dan icdrate sebesar 0,42. Kebaikan hasil klaster (R^2) sebesar 58%. Setiap klaster memiliki karakteristik masing - masing. Klaster 1 yang beranggotakan Provinsi DKI Jakarta merupakan klaster yang memiliki Indeks Kesehatan Keluarga tertinggi. Klaster 2 termasuk dalam kategori yang memiliki Indeks Kesehatan Keluarga sedang/cukup. Klaster 3 termasuk dalam kategori rendah. Sedangkan klaster 4 merupakan klaster yang memiliki Indeks Kesehatan Keluarga terendah yang memiliki anggota provinsi Papua dan Papua Barat. Terjadi kesenjangan atau perbedaan pembangunan Kesehatan di wilayah Indonesia.

Kata kunci: Analisis Klaster; Keluarga Sehat; Icdrate; Pseudo-F Statistics.

PENDAHULUAN

Pembangunan kesehatan merupakan isu penting dan fundamental yang harus segera dilakukan oleh negara Indonesia. Berdasarkan laporan *The Legatum Prosperity Index* 2017, indeks kesehatan global Indonesia berada di posisi ke 101 dari 149 negara. Kesehatan menjadi tujuan ketiga dalam *Sustainable Development Goals* (SDGs). Pembangunan kesehatan harus dimulai dari unit terkecil masyarakat, yaitu keluarga. Pemerintah pusat dan pemerintah daerah menetapkan kebijakan pembangunan keluarga melalui pembinaan ketahanan dan kesejahteraan keluarga. Hal ini bertujuan agar keluarga dapat melaksanakan fungsinya secara optimal. Kementerian kesehatan menetapkan strategi operasional pembangunan kesehatan melalui Program Indonesia Sehat (PIS) dengan Pendekatan Keluarga (PK).

Program Indonesia Sehat dengan Pendekatan Keluarga merupakan program yang diselenggarakan oleh Kemenkes RI untuk mewujudkan masyarakat Indonesia yang berperilaku sehat, hidup dalam lingkungan yang sehat, serta sadar akan pentingnya kesehatan. Sasaran dari Program Indonesia Sehat adalah meningkatnya derajat kesehatan dan status gizi masyarakat melalui upaya kesehatan dan pemberdayaan masyarakat yang

didukung dengan perlindungan finansial dan pemerataan pelayanan kesehatan (Kemenkes RI, 2016). Dalam rangka pelaksanaan Program Indonesia Sehat telah disepakati adanya 12 indikator utama untuk penanda status kesehatan sebuah keluarga berdasarkan peraturan pemerintah tahun 2016, yaitu (1) Keluarga mengikuti program Keluarga Berencana (KB); (2) Ibu melakukan persalinan di fasilitas kesehatan; (3) Bayi mendapat imunisasi dasar lengkap; (4) Bayi mendapat Air Susu Ibu (ASI) eksklusif; (5) Balita mendapatkan pemantauan pertumbuhan; (6) Penderita tuberkulosis paru mendapatkan pengobatan sesuai standar; (7) Penderita hipertensi melakukan pengobatan secara teratur; (8) Penderita gangguan jiwa mendapatkan pengobatan dan tidak ditelantarkan; (9) Anggota keluarga tidak ada yang merokok; (10) Keluarga sudah menjadi anggota JKN; (11) Keluarga mempunyai akses sarana air bersih; (12) Keluarga mempunyai akses atau menggunakan jamban sehat.

Saat ini masih terdapat kesenjangan dalam bidang kesehatan di Indonesia. Kesenjangan dapat terjadi karena perbedaan daerah, fasilitas kesehatan dan juga akses kesehatan. Oleh karena itu, diperlukan pengelompokan provinsi di Indonesia berdasarkan indikator keluarga sehat. Tujuan penelitian ini adalah untuk mengelompokkan provinsi di Indonesia berdasarkan indikator keluarga sehat. Pengelompokan dilakukan untuk mengetahui provinsi-provinsi mana yang memiliki kualitas kesehatan rendah sehingga diperlukan perbaikan akses dan fasilitas kesehatan yang lebih memadai. Pada penelitian ini dilakukan perbandingan metode klaster hierarki dan non hierarki, kemudian dengan metode terbaik tersebut yang dipakai dalam pengelompokan provinsi. Pemilihan metode terbaik berdasarkan nilai *Pseudo-F statistics* dan *icdrate*.

METODE

A. Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Profil Kesehatan Indonesia (Kementerian Kesehatan) tahun 2019 dan Statistik Indonesia (Badan Pusat Statistik) tahun 2020. Data yang digunakan merupakan data Indikator Keluarga Sehat di 34 Provinsi di Indonesia pada tahun 2019.

Variabel yang digunakan sebanyak 10 variabel dan memiliki tipe data numerik dengan skala data rasio. Variabel yang digunakan dalam penelitian ini merupakan variabel – variabel yang termasuk dalam Indikator Keluarga Sehat. Berdasarkan Peraturan Menteri Kesehatan Republik Indonesia Nomor 39 Tahun 2016 Tentang Pedoman Penyelenggaraan Program Indonesia Sehat, terdapat 12 indikator utama sebagai penanda status kesehatan sebuah keluarga. Dua belas indikator tersebut meliputi 5 indikator program gizi kesehatan ibu anak, 3 indikator pengendalian penyakit menular dan tidak menular dan 3 indikator perilaku dan kesehatan lingkungan. Namun pada penelitian ini hanya menganalisis 10 variabel karena keterbatasan data sekunder. Kesepuluh variabel ini yang paling mendekati Indikator Keluarga Sehat dan sudah merepresentasikan 12 indikator utama keluarga sehat. Variabel yang akan dianalisis terdapat pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan
X ₁	Persentase peserta KB aktif
X ₂	Persentase Ibu yang melakukan persalinan di fasilitas kesehatan
X ₃	Persentase imunisasi dasar lengkap pada bayi
X ₄	Persentase bayi yang mendapat ASI eksklusif
X ₅	Persentase balita yang pernah mendapat imunisasi dasar lengkap
X ₆	Angka keberhasilan pengobatan TBC
X ₇	Cakupan kepesertaan Jaminan Kesehatan Nasional (JKN)
X ₈	Persentase rumah tangga yang memiliki akses terhadap sumber air minum layak
X ₉	Persentase rumah tangga yang memiliki fasilitas tempat buang air besar sendiri
X ₁₀	Persentase desa yang melaksanakan sanitasi total berbasis masyarakat

B. Metode

Metode dalam penelitian yaitu metode kuantitatif dengan analisis data menggunakan metode analisis kluster. Analisis kluster adalah salah satu dari metode dalam analisis multivariat yang memiliki tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis kluster mengelompokkan individu atau objek penelitian, sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam kluster yang sama (Johnson & Wichern, 2014). Tujuan analisis kluster adalah mengelompokkan n objek ke dalam c kluster, dimana objek di dalam satu kluster memiliki tingkat homogenitas yang tinggi dan antar kelompok memiliki tingkat heterogenitas yang tinggi (Nurjannah, Farmadi, & Indriani, 2014).

Analisis kluster terbagi menjadi dua, yaitu metode hierarki dan non hierarki (Hair, Black, Babin, & Anderson, 2010). Metode kluster hierarki diantaranya yaitu *Single Linkage*, *Complete Linkage*, *Average Linkage* dan *Ward*. Sedangkan metode kluster non-hierarki yang digunakan yaitu *k-means*.

Single Linkage merupakan metode untuk mengelompokkan objek ke dalam kluster berdasarkan kedekatan dengan objek lainnya. Berdasarkan prosedur pembentukan kluster di atas, untuk metode *Single Linkage* menggunakan kriteria berikut

$$d_{(UV)W} = \min\{d_{UW}, d_{VW}\} \quad (1)$$

dengan jarak antara kluster (UV) dan kluster W merupakan jarak minimal antara kluster U dan W atau antara kluster V dan W .

Complete Linkage merupakan metode untuk mengelompokkan objek ke dalam kluster berdasarkan ketidakmiripan dengan objek lainnya. Berdasarkan prosedur pembentukan *kluster* di atas, untuk metode *complete linkage* menggunakan kriteria berikut.

$$d_{(UV)W} = \max\{d_{UW}, d_{VW}\} \quad (2)$$

dengan jarak antara kluster (UV) dan kluster W merupakan jarak paling jauh antara kluster U dan W atau antara kluster V dan W .

Metode *Ward linkage* termasuk salah satu metode hierarki yang digunakan untuk meminimalkan hilangnya informasi pada dua objek yang bergabung.

Metode *k-means* merupakan metode untuk mengklasifikasikan atau mengelompokkan objek-objek (data) ke dalam K-grup (*Cluster*) berdasarkan atribut tertentu. Pengelompokan data dilakukan dengan memperhitungkan jarak terdekat antara data-data (objek observasi) dengan pusat *Cluster (centroid)*. Prinsip utama dari metode ini adalah menyusun K buah *centroid* atau rata-rata (*mean*) dari sekumpulan data berdimensi N, dimana metode ini mensyaratkan nilai K sudah diketahui sebelumnya (apriori) (Johnson & Wichern, 2014).

Jumlah kelompok optimum dapat diketahui dengan kriteria nilai *Pseudo F-statistics*. Berikut adalah persamaan yang digunakan untuk menghitung nilai *Pseudo F-statistics*:

$$Pseudo - F Statistics = \frac{\left(\frac{R^2}{k-1}\right)}{\left(\frac{1-R^2}{n-k}\right)} \quad (3)$$

dengan,

$$R^2 = \frac{(SST-SSW)}{SST} \quad (4)$$

$$SST = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (x_{ijk} - \bar{x}_j)^2 \quad (5)$$

$$SSW = \sum_{i=1}^n \sum_{j=1}^c \sum_{k=1}^p (x_{ijk} - \bar{x}_{jk})^2 \quad (6)$$

Keterangan :

SST (*Sum Square Total*) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata keseluruhan

SSW (*Sum Square Within*) : Total jumlah dari kuadrat jarak sampel terhadap rata-rata kelompoknya

n : banyaknya sampel

c : banyaknya variabel

p : banyaknya kelompok

x_{ijk} : sampel ke-*i* pada variabel ke-*j* kelompok ke-*k*

\bar{x}_j : rata-rata seluruh sampel pada variabel ke-*j*

\bar{x}_{jk} : rata-rata sampel pada variabel ke-*j* dan kelompok ke-*k*

Nilai *Pseudo F-statistics* tertinggi menunjukkan bahwa jumlah kelompok yang digunakan untuk mempartisi data secara optimal, dimana keragaman dalam kelompok sangat homogen sedangkan antar kelompok sangat heterogen (Orpin & Kostylev, 2006).

Salah satu metode yang digunakan untuk memilih metode klaster terbaik adalah menghitung rata-rata persebaran *Internal Cluster Dispersion Rate (icdrate)* terhadap partisi keseluruhan. *Icdrate (internal cluster dispersion)* menggambarkan tingkat dispersi atau perbedaan dalam klaster. Formula untuk menghitung *icdrate* terdapat di persamaan (6). Semakin kecil nilai *icdrate* maka semakin baik hasil pengelompokannya (Mingoti & Lima, 2006).

$$icdrate = 1 - \frac{SSB}{SST} = 1 - \frac{SST-SSW}{SST} = 1 - R^2 \quad (6)$$

Keterangan :

SSB : *Sum Square Between* (SST-SSW)

R^2 : *Recovery Rate* (SSB/SST)

HASIL DAN PEMBAHASAN

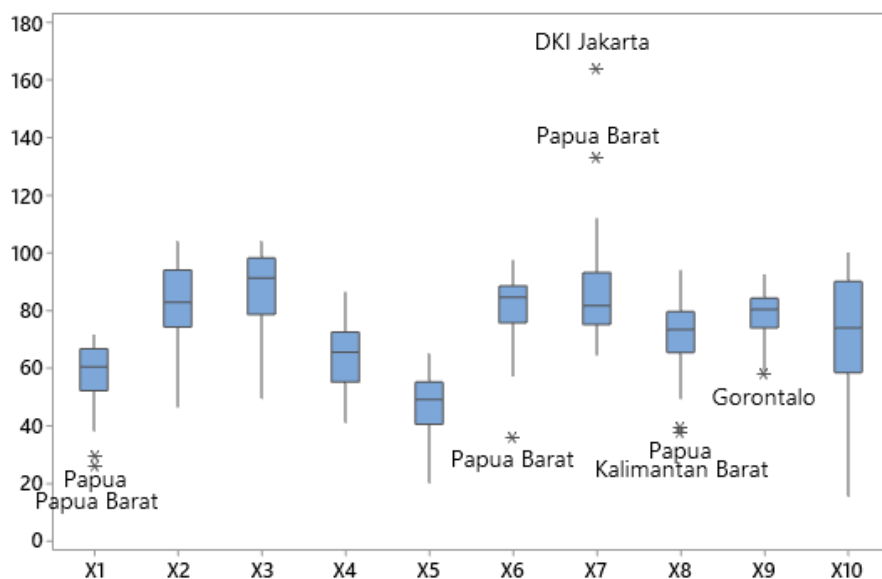
A. Gambaran Umum Data Indikator Keluarga Sehat

Karakteristik dan gambaran dari data Indikator Keluarga Sehat perlu diketahui sebelum melakukan analisis. Untuk mengetahui gambaran umum tentang data persentase Indikator Keluarga Sehat di Indonesia berdasarkan Provinsi, dilakukan analisis statistika deskriptif. Tabel 2 menunjukkan hasil analisis statistika deskriptif. Berdasarkan Tabel 2 dapat diketahui bahwa masing-masing Indikator Keluarga Sehat memiliki rata – rata yang berbeda-beda. Rata – rata yang terendah yaitu persentase balita yang pernah mendapat imunisasi dasar lengkap (X_5) yakni sebesar 47,13%. Hal tersebut menunjukkan bahwa masih sedikit balita di Indonesia yang mendapat imunisasi dasar lengkap. Sedangkan, rata – rata yang tertinggi yaitu cakupan kepesertaan Jaminan Kesehatan Nasional (X_7). Hal ini menunjukkan bahwa cakupan masyarakat dalam kepesertaan Jaminan Kesehatan Nasional (JKN) sudah baik. Nilai minimum merupakan nilai paling kecil dari keseluruhan data di setiap provinsi untuk masing – masing variabel. Nilai maksimum merupakan nilai yang paling besar dari keseluruhan data di setiap provinsi untuk masing – masing variabel.

Tabel 2. Statistika Deskriptif

Variabel	Mean	Variance	Minimum	Maksimum
X_1	57,68	131,88	25,4	71,4
X_2	81,28	202,17	46,60	103,8
X_3	87,42	182,10	49,60	103,8
X_4	64,05	133,33	41,12	86,26
X_5	47,13	114,01	20,18	64,79
X_6	81,12	147,50	35,60	97,30
X_7	86,98	367,01	64,68	164,32
X_8	72,14	161,05	37,37	93,77
X_9	78,99	77,50	57,65	92,49
X_{10}	71,38	488,47	15,45	100

Tabel 2 menunjukkan bahwa persentase rumah tangga yang memiliki fasilitas tempat buang air besar sendiri (X_9) memiliki varians atau ragam yang paling kecil yaitu sebesar 77,50, sehingga dapat dikatakan bahwa masing – masing provinsi di Indonesia memiliki Persentase rumah tangga yang memiliki fasilitas tempat buang air besar sendiri yang hampir sama atau tersebar secara merata. Namun untuk variabel lain memiliki varians yang tinggi (lebih besar dari 100) sehingga persebarannya tidak merata. Varians merupakan ukuran variabilitas data, yang berarti semakin besar nilai varians berarti semakin tinggi fluktuasi data antara satu data dengan data yang lain.



Gambar 1. *Boxplot* Setiap Variabel Indikator Keluarga Sehat

Visualisasi gambaran umum dan distribusi data dapat dilihat pada Gambar 2. Gambar 1 (*boxplot*) terdapat garis tengah pada kotak menunjukkan nilai median. Semakin panjang *boxplot* menunjukkan bahwa data semakin menyebar atau semakin bervariasi. Variabel Persentase desa yang melaksanakan sanitasi total berbasis masyarakat (X_{10}) memiliki panjang *boxplot* yang paling panjang, artinya variabel tersebut memiliki ragam yang paling tinggi. Terlihat bahwa dari beberapa variabel terdapat *outlier* atau pencilan. *Outlier* adalah data yang menyimpang terlalu jauh dari data yang lainnya atau dengan kata lain memiliki nilai ekstrem. Pada variabel Persentase peserta KB aktif (X_1) provinsi yang termasuk kategori *outlier* yaitu Papua dan Papua Barat. Variabel X_2 , X_3 , X_4 , X_5 , dan X_{10} tidak ada data yang *outlier*. Variabel Angka keberhasilan pengobatan TBC (X_6) terdapat *outlier* di provinsi Papua Barat. Variabel cakupan kepesertaan Jaminan Kesehatan Nasional (X_7) memiliki *outlier* pada provinsi Papua Barat dan DKI Jakarta. Pada variabel persentase rumah tangga yang memiliki akses terhadap sumber air minum layak (X_8) provinsi yang *outlier* yaitu Papua dan Kalimantan Barat. Variabel persentase rumah tangga yang memiliki fasilitas tempat buang air besar sendiri (X_9) memiliki *outlier* pada provinsi Gorontalo.

B. Hasil Analisis Kluster Hierarki dan Non-Hierarki Indikator Keluarga Sehat

Pada penelitian pengelompokan menggunakan metode kluster hirarki dan non-hirarki masing-masing akan dibentuk menjadi 3, 4 dan 5 kluster. Pembentukan menjadi 3, 4, dan 5 kluster ini berdasarkan kelompok indikator yang akan dibentuk. Jika kluster yang optimum adalah 3, maka pengelompokan indikator keluarga sehat menjadi kategori rendah, sedang, dan tinggi. Jika kluster optimum yang terbentuk adalah 4, maka kategori indikator keluarga sehat menjadi sangat rendah, rendah, sedang, dan tinggi. Jika kluster optimum yang terbentuk sebanyak 5 kluster, maka kategori indikator keluarga sehat yang terbentuk yaitu sangat rendah, rendah, sedang, tinggi, dan sangat tinggi.

Pemilihan banyak kluster yang optimum menggunakan *Pseudo-F Statistics*. Perbandingan nilai *Pseudo-F Statistics* metode kluster hirarki dan non-hirarki dengan banyak kelompok 3,4 dan 5 ditampilkan pada Tabel 3.

Tabel 3. *Pseudo-F Statistics*

Jumlah Kluster	<i>Single Linkage</i>	<i>Complete Linkage</i>	<i>Average Linkage</i>	<i>Ward Linkage</i>	<i>K-Means</i>
3	9,63	10,28	12,33	12,96	12,72
4	6,81	11,08	8,86	13,03	13,57
5	6,05	9,59	6,43	11,65	11,33

Berdasarkan Tabel 3, Nilai *Pseudo-F Statistics* terbesar dari semua kombinasi metode dan jumlah kluster adalah sebesar 13,57 yang terdapat pada kelompok dengan 4 kluster dengan menggunakan metode *K-means*. Nilai *Pseudo F-statistics* yang tinggi menunjukkan bahwa dengan 4 kluster dapat menghasilkan keragaman dalam kelompok sangat homogen sedangkan antar kelompok sangat heterogen. Hal ini menunjukkan bahwa banyak kluster optimum yang terbentuk adalah 4 kluster. Kategori 4 kluster yang terbentuk yaitu indikator keluarga sehat sangat rendah, rendah, sedang, dan tinggi. Kemudian dengan jumlah 4 kluster dilakukan pengelompokan menggunakan metode *Single Linkage*, *Complete Linkage*, *Average Linkage*, *Ward Linkage*, dan *K-Means*. Lalu dihitung nilai *icdrate* untuk menentukan metode yang terbaik. Metode yang terbaik memiliki nilai *icdrate* terkecil dan R^2 terbesar. Hasil perbandingan *icdrate* dan R^2 ditampilkan pada Tabel 4.

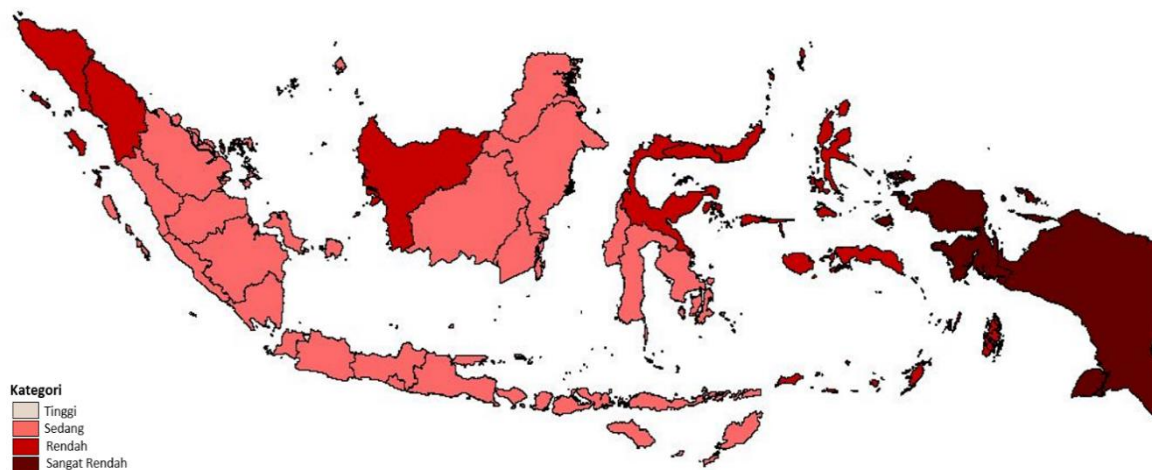
Tabel 4. *Icdrate* dan R^2

	<i>Single Linkage</i>	<i>Complete Linkage</i>	<i>Average Linkage</i>	<i>Ward Linkage</i>	<i>K-Means</i>
<i>Icdrate</i>	0,62	0,47	0,56	0,43	0,42
R^2	0,38	0,53	0,44	0,54	0,58

Metode *k-means* memiliki *icdrate* terkecil yaitu sebesar 0,42 dan R^2 terbesar yaitu sebesar 0,58. Dispersi atau perbedaan dalam kluster sebesar 42% dan kebaikan hasil kluster sebesar 58%. Meskipun hasil kebaikan kluster tidak terlalu baik, namun kebaikan kluster ini sudah diatas 50%, sehingga cukup reliabel untuk diinterpretasikan lebih lanjut. Semakin kecil nilai *icdrate* maka semakin baik hasil pengelompokannya. Sehingga metode terbaik dalam mengelompokkan provinsi di Indonesia berdasarkan Indikator Keluarga Sehat dengan jumlah pengelompokan optimal adalah 4 kluster adalah metode non-hirarki (*k-means*). Dengan menggunakan algoritma atau metode *k-means* didapatkan anggota pada masing – masing kluster yang disajikan pada Tabel 5. Pembagian anggota pada masing – masing kluster berdasarkan kedekatan (*similarity*). Visualisasi atau peta pengelompokan provinsi di Indonesia berdasarkan indikator keluarga sehat disajikan dalam Gambar 2.

Tabel 5. Rincian Anggota Klaster

Nomor Klaster	Klaster	Banyak Anggota	Anggota dalam Klaster
1	Tinggi	1	DKI Jakarta
2	Sedang/Cukup	23	Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kep. Riau, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Selatan, Sulawesi Tenggara, Sulawesi Barat.
3	Rendah	8	Aceh, Sumatera Utara, Kalimantan Barat, Sulawesi Utara, Sulawesi Tengah, Gorontalo, Maluku, Maluku Utara
4	Sangat Rendah	2	Papua, Papua Barat



Gambar 2. Peta Hasil Pengelompokan dengan Metode Klaster Terbaik

Terlihat pada peta di Gambar 2 bahwa wilayah Indonesia bagian timur, khususnya Papua dan Papua barat memiliki Indeks Kesehatan Keluarga yang sangat rendah. Wilayah pulau Sumatera bagian utara, Kalimantan barat, dan Sulawesi Tengah daerah kepulauan maluku memiliki Indeks Kesehatan Keluarga. Sedangkan di wilayah, Jawa, Bali, Sumatera bagian tengah hingga selatan, Pulau Kalimantan Tengah hingga Timur, Sebagian pulau Sulawesi sudah memiliki Indeks Kesehatan Keluarga yang cukup. Terlihat bahwa terjadi kesenjangan Indeks Kesehatan Keluarga di wilayah Indonesia.

Tabel 6. Rata – rata Indikator Keluarga Sehat Setiap Klaster

Variabel	Rata – rata			
	Klaster 1 (Tinggi)	Klaster 2 (Sedang/Cukup)	Klaster 3 (Rendah)	Klaster 4 (Sangat Rendah)
X ₁	56,90	60,86	56,23	27,25
X ₂	103,80	85,19	75,10	49,65
X ₃	98,00	91,67	78,31	69,85
X ₄	70,22	69,46	53,44	41,27
X ₅	52,09	49,96	41,10	36,105
X ₆	77,60	85,06	77,60	51,75
X ₇	164,32	80,92	85,83	122,69
X ₈	88,44	73,76	69,83	54,62
X ₉	83,02	81,95	73,23	65,86
X ₁₀	94,01	81,26	53,34	18,64
Rata – rata total	88,84	76,01	66,40	53,77

Karakteristik setiap klaster yang terbentuk dengan metode Non-Hierarki (*k-means*) terdapat pada Tabel 6. Klaster 1 memiliki nilai X₂, X₃, X₄, X₅, X₇, X₈, X₉, dan X₁₀ paling tinggi di antara klaster yang lain. Hasil rata – rata total menunjukkan bahwa klaster 1 mempunyai rata – rata indeks kesehatan keluarga tertinggi.

Klaster 2 memiliki karakteristik variabel yang nilainya bervariasi, dari yang rendah sampai tertinggi. Klaster 2 memiliki 2 Indikator Keluarga Sehat yang paling tinggi dari pada klaster lainnya yaitu Persentase peserta KB aktif (X₁) dan Angka keberhasilan pengobatan TBC (X₆). Klaster ini juga memiliki 1 Indikator Keluarga Sehat yang paling rendah yaitu Cakupan kepesertaan Jaminan Kesehatan Nasional (X₇). Secara keseluruhan klaster 2 merupakan klaster dengan Indikator Keluarga Sehat sedang.

Secara keseluruhan klaster 3 menjadi klaster dengan Indikator Keluarga Sehat yang rendah dengan rata – rata total sebesar 66,40. Pada klaster ini tidak ada variabel dari Indikator Keluarga Sehat yang tertinggi maupun terendah. Namun, nilai variabel Indikator Keluarga Sehat secara keseluruhan tidak terlalu tinggi, sehingga dapat dikatakan klaster 3 menjadi klaster rendah.

Sedangkan klaster 4 memiliki nilai X₁, X₂, X₃, X₄, X₅, X₆, X₈, X₉, dan X₁₀ terendah diantara klaster lain. Hasil rata – rata total menunjukkan bahwa klaster 4 mempunyai rata – rata indeks Kesehatan keluarga terendah. Pengelompokan provinsi dengan kategori Indikator Keluarga Sehat sangat rendah, rendah, sedang dan tinggi dapat dijelaskan pada Gambar 2.

KESIMPULAN

Pengelompokan provinsi di Indonesia berdasarkan Indikator Keluarga Sehat dengan analisis klaster hierarki dan non-hierarki menghasilkan 4 klaster. Metode terbaik antara klaster hierarki dan non hierarki yaitu metode *k-means* dengan 4 klaster dihasilkan nilai *Pseudo-F statistics* sebesar 13,57 dan *icdrate* sebesar 0,42. Keباikan hasil klaster (R^2) sebesar 58%. Setiap klaster memiliki karakteristik masing – masing. Klaster 1 yang beranggotakan Provinsi DKI Jakarta merupakan klaster yang memiliki Indeks Kesehatan Keluarga tertinggi. Klaster 2 termasuk dalam kategori yang memiliki Indeks Kesehatan Keluarga sedang atau cukup. Klaster 3 termasuk dalam kategori rendah. Sedangkan

klaster 4 merupakan klaster yang memiliki Indeks Kesehatan Keluarga terendah dengan anggota klasternya yaitu provinsi Papua dan Papua Barat. Terjadi kesenjangan atau perbedaan pembangunan Kesehatan di wilayah Indonesia.

UCAPAN TERIMAKASIH

Ucapan terima kasih disampaikan kepada Departemen Statistika ITS yang sudah membekali ilmu dan juga kepada Badan Pusat Statistik (BPS) yang sudah mempublikasi data.

DAFTAR PUSTAKA

- BPS. (2020). *Statistik Indonesia*. Jakarta: Badan Pusat Statistik.
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate Data Analysis, Seventh Edition*. England: Pearson.
- Johnson, R., & Wichern, D. (2014). *Applied Multivariate Statistical Analysis (Sixth Edition)*. Edinburg: Pearson.
- Kemendes RI. (2016). *Pedoman Umum Program Indonesia Sehat dengan Pendekatan Keluarga*. Jakarta: Kemendes RI.
- Mingoti, S., & Lima, J. (2006). Comparing SOM Neural Network with Fuzy C-Means, K-Means and Traditional Hierarchical Clustering Algorithm. *European Journal of Operational Research*, 1742-1759.
- Nurjannah, Farmadi, A., & Indriani, F. (2014). Implementasi Metode Fuzzy C-Means Pada Sistem Clustering Data Varietas Padi. *Kumpulan jurnal Ilmu Komputer (KLIK)*, 23-32.
- Orpin, A., & Kostylev, V. (2006). Toward a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats. *Marine Geology*, 209-222.