

# Pengelompokan Kota/Kabupaten di Provinsi Jawa Barat Berdasarkan Indikator Keluarga Sehat Tahun 2019 Menggunakan Metode *Average Linkage*

Sophia Fauziah\*, Lisnur Wachidah

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

\*sophiafzh02@gmail.com, wachidah.lisnur07@gmail.com

**Abstract.** Cluster analysis is a method to form several subgroups by grouping objects that have characteristics according to their characteristics. Cluster analysis is divided into two methods, namely hierarchical and non-hierarchical. The hierarchical method is used to group objects with relatively small data sizes where the resulting cluster will be displayed on the dendrogram. The hierarchical way has two types of approaches, namely agglomerative and divisional. The agglomerative process starts with the individual as the object. The objects that have ties will first merge into sub-groups, and eventually, these sub-groups will merge into one cluster. There are five agglomerative approaches, namely single linkage, complete linkage, average linkage, ward, and centroid. The weakness of the single linkage, complete linkage, ward, and centroid methods is that they cannot estimate every distance between points and cannot observe the average distance between all cluster observations. Therefore, this study uses average linkage because it can consider each distance between points and the average distance between all cluster observations. For this problem, secondary data will be used, namely healthy family indicator data in 2019. The results of the grouping of 27 cities/districts in West Java Province resulted in 4 clusters, namely cluster 1 consisting of 8 city/district members with an average family health condition of 39.79%, cluster 2 consists of 12 city/district members with an average family health condition score of 31.61%, cluster 3 consists of 5 city/district members with an average family health condition score of 40.44% and cluster 3 4 consists of 2 city/district members with an average family health condition of 49.65%.

**Keywords:** Average Linkage, Cluster, Hierarki Method, Healthy Family Indicator.

**Abstrak.** Analisis *cluster* adalah suatu metode untuk membentuk beberapa subkelompok dengan mengelompokkan objek-objek yang mempunyai kemiripan menurut karakteristiknya. Analisis *cluster* terbagi menjadi 2 metode yaitu hierarki dan non-hierarki. Metode hierarki adalah metode yang digunakan untuk mengelompokkan objek dengan ukuran data yang relatif sedikit dimana *cluster* yang dihasilkan akan ditampilkan pada dendrogram. Metode hierarki memiliki dua jenis pendekatan yaitu *agglomerative* dan *divisive*. Pendekatan *agglomerative* dimulai dengan individu sebagai objek kemudian objek yang memiliki kemiripan akan digabungkan terlebih dahulu menjadi sub kelompok, pada akhirnya sub kelompok ini akan digabung menjadi satu *cluster*. Terdapat lima metode pada pendekatan *agglomerative* yaitu *single linkage*, *complete linkage*, *average linkage*, *ward* dan *centroid*. Kelemahan dari metode *single linkage*, *complete linkage*, *ward* dan *centroid* adalah tidak dapat memperhitungkan setiap jarak antar titiknya dan tidak dapat meminimalkan rata-rata jarak antar semua pengamatan *cluster*. Maka dari itu, pada penelitian ini menggunakan *average linkage* karena dapat memperhitungkan setiap jarak antar titiknya dan meminimalkan rata-rata jarak antara semua pengamatan pasangan *cluster*. Untuk masalah ini akan digunakan pada data sekunder yaitu data indikator keluarga sehat tahun 2019. Hasil

pengelompokan dari 27 kota/kabupaten di Provinsi Jawa Barat menghasilkan 4 *cluster* yaitu *cluster* 1 terdiri dari 8 anggota kota/kabupaten dengan nilai rata-rata kondisi kesehatan keluarga sebesar 39,79%, *cluster* 2 terdiri dari 12 anggota kota/kabupaten dengan nilai rata-rata kondisi kesehatan keluarga sebesar 31,61% *cluster* 3 terdiri dari 5 anggota kota/kabupaten dengan nilai rata-rata kondisi kesehatan keluarga sebesar 40,44% dan *cluster* 4 terdiri dari 2 anggota kota/kabupaten dengan nilai rata-rata kondisi kesehatan keluarga sebesar 49,65%.

**Kata Kunci:** *Average Linkage, Cluster, Metode Hierarki, Indikator Keluarga Sehat.*

## 1. Pendahuluan

Program Indonesia Sehat merupakan program utama pemerintah dalam mencapai pembangunan kesehatan tahun 2015 sampai 2019 yang bertujuan untuk meningkatkan kualitas hidup manusia Indonesia (Kementerian Kesehatan RI, 2016:1). Dalam mencapai pembangunan kesehatan Indonesia diperlukan dukungan dari berbagai pihak, dimana keluarga merupakan bagian dari unit terkecil yang turut berperan penting untuk mendukung pembangunan kesehatan (Kementerian Kesehatan RI, 2016:19).

Berdasarkan laporan data penelitian Dinas Kesehatan Provinsi Jawa Barat tahun 2019 menunjukkan terjadi penurunan cakupan persentase indikator keluarga sehat di tahun 2019. diantaranya yaitu: (1) keluarga mengikuti program KB (2) penderita TBC Paru yang berobat sesuai standar (3) penderita hipertensi yang berobat secara teratur (4) penderita gangguan jiwa berat diobat tidak terlantar (5) anggota keluarga tidak ada yang merokok (6) keluarga sudah menjadi anggota jaminan kesehatan nasional.

Status kesehatan keluarga Jawa Barat yang tidak seragam akan menyebabkan pemerintah daerah mengalami kesulitan dalam melakukan pembinaan. Provinsi Jawa Barat terdiri dari 27 kota/kabupaten yang masing-masing memiliki kondisi kesehatan keluarga yang berbeda. Pengelompokan wilayah kota/kabupaten di Jawa Barat perlu dilakukan berdasarkan kesamaan karakteristik kondisi kesehatan keluarga berdasarkan faktor-faktor dengan menggunakan metode average linkage.

Pada metode hierarki terdapat dua cara untuk mengelompokkan yaitu agglomerative dan divisive (Hanada Silvia, 2021). Metode average linkage adalah metode mencari nilai kesamaan dengan menghitung jarak rata-rata diantara semua pasangan titik milik anggota dari setiap pasangan titik milik cluster dimana menggabungkan jarak dua cluster dengan rata-rata jarak terkecil. Sedangkan jika menggunakan metode lain maka dapat menyebabkan penggabungan yang kurang optimal jika mendapati data yang digunakan mengandung pencilan (Aggarwal dan Reddy, 2013). Metode average linkage memberikan hasil pengelompokan, karena metode average linkage dapat memperhitungkan setiap jarak antar titiknya dan meminimalkan rata-rata jarak antara semua pengamatan pasangan cluster ("Scikit Learn", n.d.).

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini adalah "Bagaimana menerapkan metode average linkage dalam pengelompokan kota/kabupaten di Provinsi Jawa Barat berdasarkan indikator keluarga sehat tahun 2019". Selanjutnya, tujuan dalam penelitian ini yaitu menerapkan metode average linkage untuk mengidentifikasi jenis penyakit dan nilai rata-rata kondisi kesehatan keluarga dalam pengelompokan kota/kabupaten berdasarkan indikator keluarga sehat di Provinsi Jawa Barat Tahun 2019.

## 2. Metodologi

### Bahan

Jenis data yang digunakan pada penelitian ini yaitu data sekunder hasil laporan yang diperoleh dari Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia. Data tersebut berisi

rekapitulasi indikator keluarga sehat untuk semua wilayah di Indonesia, kemudian data yang akan digunakan pada penelitian ini adalah indikator keluarga sehat per kota/kabupaten pada wilayah Jawa Barat tahun 2019 dengan sampel pengamatan sejumlah 27 kota/kabupaten di Jawa Barat.

### Variabel Penelitian

Variabel yang digunakan adalah persentase Indikator keluarga sehat yang dapat dilihat pada Tabel 1 di bawah ini:

**Tabel 1.** Variabel Penelitian

### Metode Analisis Penelitian

1. Analisis Deskriptif dan eksplorasi data untuk setiap masing-masing variabel indikator keluarga sehat
2. Masukkan data persentase keluarga mengikuti program KB, penderita TBC Paru yang berobat sesuai standar, penderita gangguan jiwa berat diobati tidak terlantar, anggota keluarga tidak ada yang merokok, keluarga sudah menjadi anggota jaminan kesehatan nasional, dan penderita hipertensi yang berobat secara teratur.
3. Menguji kecukupan sampel apakah sampel yang digunakan dapat mewakili populasi dengan menggunakan uji KMO (Kaiser Meyer Olkin)

Variabel	Atribut	Skala	Satuan
$X_1$	Keluarga Mengikuti Program KB	Rasio	%
$X_2$	Penderita TBC Paru yang Berobat Sesuai Standar	Rasio	%
$X_3$	Penderita Hipertensi yang Berobat Secara Teratur	Rasio	%
$X_4$	Penderita Gangguan Jiwa Diobati dan Tidak Terlantar	Rasio	%
$X_5$	Anggota Keluarga Tidak Ada yang Merokok	Rasio	%
$X_6$	Keluarga Sudah Menjadi Anggota JKN	Rasio	%

4. Melakukan uji multikolinearitas untuk melihat terjadi hubungan antar variabel atau tidak pada data
5. Melakukan pengelompokkan kota/kabupaten menggunakan metode average linkage menggunakan langkah-langkah sebagai berikut:
  - a. Mengukur kemiripan antar objek (dua kota/kabupaten) menggunakan rumus jarak euclidean
  - b. Menentukan jarak yang terdekat antar objek (kota/kabupaten).
  - c. Dengan menggunakan metode average linkage lakukan perhitungan ulang.
  - d. Menentukan jumlah *cluster* dan anggota pada *cluster* yang terbentuk
  - e. Interpretasi Cluster

## 3. Pembahasan dan Diskusi

### Analisis Deskriptif

Deskripsi data yaitu melakukan eksplorasi data yang dilakukan untuk mendapatkan gambaran umum dari data yang digunakan pada penelitian. Data penelitian ini merupakan hasil dari rekapitulasi indikator keluarga sehat pada Provinsi Jawa Barat tahun 2019, dengan sampel pengamatan sejumlah 27 kota/kabupaten di Jawa Barat.

**Tabel 2.** Statistik Deskriptif

Variabel	N	Maximum	Minimum	Rata-rata
$x_1$	27	64.01	12.26	29.61
$x_2$	27	58.57	26.14	42.65
$x_3$	27	40.95	15.79	26.10
$x_4$	27	48.36	27.54	37.15
$x_5$	27	52.68	25.22	36.27
$x_6$	27	76.70	31.41	50.25

Berdasarkan Tabel 2 rata-rata terbesar didapatkan untuk keluarga sudah menjadi anggota JKN ( $x_6$ ) adalah sebesar 29,61. Nilai tersebut menunjukkan bahwa secara rata-rata keluarga sudah menjadi anggota JKN yaitu terbesar dari seluruh kab/kota di Jawa Barat dan keluarga dengan penderita hipertensi yang berobat secara teratur yaitu terendah dari seluruh kab/kota di Jawa Barat.

Analisis *cluster hierarki* diawali dengan menguji asumsi analisis *cluster*. Asumsi analisis *cluster hierarki* yang pertama adalah uji kaiser meyer olkin (KMO) yaitu uji kecukupan data untuk dilakukan analisis. Jika nilai KMO berada diantara 0,5 hingga 1 maka sampel dikatakan cukup atau dikatakan bahwa sampel mewakili populasi (Widarjono, 2010)

**Tabel 3.** Uji KMO

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0,657
Barlett's Test of Sphericity	Approx. Chi-Square	62,033
	Df	15
	Sig.	0,000

Berdasarkan hasil KMO didapat nilai sebesar 0,657 yang artinya data indikator keluarga sehat kabupaten/kota di Provinsi Jawa Barat pada tahun 2019 mewakili populasi sehingga layak untuk dilakukan analisis *cluster*.

Uji asumsi *cluster* selanjutnya adalah uji multikolinearitas digunakan untuk melihat apakah terdapat multikolinearitas diantara variabel-variabel bebas, dengan menggunakan nilai VIF jika nilai VIF lebih besar dari 10 maka terjadinya multikolinearitas (Kurniawan dan Yuniarto, 2016).

**Tabel 4.** Uji VIF Antara Variabel  $x_1$  dengan variabel bebas  $x_2, x_3, x_4, x_5$ , dan  $x_6$

Collinearity Statistics	
Constant	VIF
$X_2$	2,007
$X_3$	1,477
$X_4$	1,696
$X_5$	3,298
$X_6$	5,013

**Tabel 5.** Uji VIF Antara Variabel  $x_2$  dengan variabel bebas  $x_3, x_4, x_5$ , dan  $x_6$

Collinearity Statistics	
Constant	VIF
$X_3$	1,477
$X_4$	1,361

$X_5$	3,248
$X_6$	4,473

**Tabel 6.** Uji VIF Antara Variabel  $x_3$  dengan variabel bebas  $x_4, x_5$ , dan  $x_6$

<i>Collinearity Statistics</i>	
Constant	VIF
$X_4$	1,346
$X_5$	3,113
$X_6$	3,694

**Tabel 7.** Uji VIF Antara Variabel  $x_4$  dengan variabel bebas  $x_5$ , dan  $x_6$

<i>Collinearity Statistics</i>	
Constant	VIF
$X_5$	2,987
$X_6$	2,987

**Tabel 8.** Uji VIF Antara Variabel  $x_5$  dengan variabel bebas  $x_6$

<i>Collinearity Statistics</i>	
Constant	VIF
$X_6$	1,000

Dapat dilihat bahwa dari Tabel 4 – 7 bahwa nilai VIF antara semua variabel bebas menunjukkan nilai  $VIF < 10$  yang artinya tidak terjadi multikolinearitas antar variabel tersebut. Sehingga disimpulkan tidak terjadi multikolinearitas antar semua variabel bebas pada data yang digunakan dalam skripsi ini.

Hasil pembagian kelompok dengan menggunakan metode *average linkage* menghasilkan 4 *cluster* dengan anggota pada masing-masing *cluster* ditampilkan pada Tabel 9.

**Tabel 9.** Anggota *Cluster*

<i>Cluster</i>	Anggota <i>Cluster</i>
<i>Cluster 1</i>	Kab. Bogor, Kab. Bandung, Kab. Ciamis, Kab. Kuningan, Kab. Cirebon, Kab. Purwakarta, Kota Sukabumi dan Kota Bandung
<i>Cluster 2</i>	Kab. Sukabumi, Kab. Cianjur, Kab. Garut, Kab. Tasikmalaya, Kab. Majalengka, Kab. Sumedang, Kab. Indramayu, Kab. Subang, Kab. Bandung Barat, Kab. Pangandaran, Kota Tasikmalaya dan Kota Banjar

<i>Cluster 3</i>	Kab. Karawang, Kab. Bekasi, Kota Bogor, Kota Bekasi dan Kota Depok
<i>Cluster 4</i>	Kota Cirebon dan Kota Cimahi

**Tabel 10.** Identifikasi Jenis Penyakit *Cluster*

Variabel	<i>Cluster</i>			
	1	2	3	4
$X_1$	32,89%	27,63%	17,57%	58,49%
$X_2$	46%	37,77%	46,30%	49,34%
$X_3$	29,20%	22,73%	29,48%	25,46%
$X_4$	42%	33,05%	36,18%	44,78%
$X_5$	35,52%	31,79%	43,83%	47,27%
$X_6$	53,12%	36,70%	69,26%	72,54%

Diketahui bahwa penyebaran status kesehatan keluarga berdasarkan indikator keluarga sehat di Provinsi Jawa Barat terbagi menjadi 4 *cluster* dengan tingkat status kesehatan keluarga yang berbeda-beda. Dapat disimpulkan sebagai berikut:

1. *Cluster 1* adalah *cluster* yang memiliki nilai rata-rata terkecil untuk indikator penderita hipertensi yang berobat secara teratur ( $X_3$ ) dan nilai rata-rata terbesar untuk indikator keluarga sudah menjadi anggota jaminan kesehatan nasional (JKN) ( $X_6$ ).
2. *Cluster 2* adalah *cluster* yang memiliki nilai rata-rata terkecil untuk indikator penderita hipertensi yang berobat secara teratur ( $X_3$ ) dan nilai rata-rata terbesar untuk indikator penderita TBC Paru yang berobat sesuai standar ( $X_2$ ).
3. *Cluster 3* adalah *cluster* yang memiliki nilai rata-rata terkecil untuk indikator keluarga mengikuti program KB ( $X_1$ ) dan nilai rata-rata terbesar untuk indikator keluarga sudah menjadi anggota jaminan kesehatan nasional (JKN) ( $X_6$ ).
4. *Cluster 4* adalah *cluster* yang memiliki nilai rata-rata terkecil untuk indikator penderita hipertensi yang berobat secara teratur ( $X_3$ ) dan nilai rata-rata terbesar untuk indikator keluarga sudah menjadi anggota jaminan kesehatan nasional (JKN) ( $X_6$ ).

Untuk nilai rata-rata kondisi kesehatan keluarga yang didapatkan untuk masing-masing *cluster*

disajikan pada Tabel 11

**Tabel 11.** Rata-Rata Kondisi Kesehatan Keluarga

<i>Cluster</i>	Rata-Rata
<i>Cluster 1</i>	39,79%
<i>Cluster 2</i>	31,61%
<i>Cluster 3</i>	40,44%
<i>Cluster 4</i>	49,65%

Dapat dilihat pada Tabel 11 yaitu *cluster 1* memiliki nilai rata-rata kondisi kesehatan keluarga sebesar 39,79%, *cluster 2* memiliki nilai rata-rata kondisi kesehatan keluarga sebesar 31,61%, *cluster 3* memiliki nilai rata-rata kondisi kesehatan keluarga sebesar 40,44%, dan *cluster 4* memiliki nilai rata-rata sebesar 49,65%. Artinya bahwa *cluster 4* memiliki nilai rata-rata kondisi kesehatan keluarga tertinggi dan *cluster 2* memiliki nilai rata-rata kondisi kesehatan keluarga terendah.

#### 4. Kesimpulan

Hasil pengelompokan 27 kota/kabupaten untuk keenam indikator keluarga sehat di Provinsi Jawa Barat tahun 2019 dengan menggunakan metode average linkage terbagi menjadi 4 cluster atau kelompok yang didapatkan dari analisis yang telah dilakukan. Adapun cluster tersebut adalah sebagai berikut:

1. Cluster 1 terdiri dari 8 anggota yaitu (1) Kabupaten Bogor, (2) Kabupaten Bandung, (3) Kabupaten Ciamis, (4) Kabupaten Kuningan, (5) Kabupaten Cirebon, (6) Kabupaten Purwakarta, (7) Kota Sukabumi dan (8) Kota Bandung. Cluster ini memiliki permasalahan jenis penyakit dengan daerah penderita hipertensi yang tidak berobat secara teratur paling banyak. Cluster 1 memiliki nilai rata-rata kondisi kesehatan keluarga sebesar 39,79%
2. Cluster 2 terdiri dari 12 anggota yaitu (1) Kabupaten Sukabumi, (2) Kabupaten Cianjur, (3) Kabupaten Garut, (4) Kabupaten Tasikmalaya, (5) Kabupaten Majalengka, (6) Kabupaten Sumedang, (7) Kabupaten Indramayu, (8) Kabupaten Subang, (9) Kabupaten Bandung Barat, (10) Kabupaten Pangandaran, (11) Kota Tasikmalaya, dan (12) Kota Banjar. Cluster ini memiliki permasalahan jenis penyakit dengan daerah penderita hipertensi yang tidak berobat secara teratur paling banyak. Cluster 2 memiliki nilai rata-rata kondisi kesehatan keluarga sebesar 31,61%
3. Cluster 3 terdiri dari 5 anggota yaitu (1) Kabupaten Karawang, (2) Kabupaten Bekasi, (3) Kota Bogor, (4) Kota Bekasi, dan (5) Kota Depok. Cluster ini memiliki permasalahan jenis penyakit dengan daerah sedikitnya keluarga yang mengikuti Program KB. Cluster 3 memiliki nilai rata-rata kondisi kesehatan keluarga sebesar 40,44%
4. Cluster 4 terdiri dari 2 anggota yaitu (1) Kota Cirebon dan (2) Kota Cimahi. Cluster ini memiliki permasalahan jenis penyakit dengan daerah penderita hipertensi yang tidak berobat secara teratur cukup banyak. Cluster 4 memiliki nilai rata-rata kondisi kesehatan keluarga sebesar 49,65%

#### Acknowledge

Terima kasih kepada seluruh pihak yang membantu memberi saran maupun masukan sampai terlaksananya penelitian ini

**Daftar Pustaka**

- [1] Kementerian Kesehatan RI. (2016). Pedoman Indonesia Sehat Dengan Pendekatan Keluarga. Jakarta: Kemenkes RI
- [2] Aggarwal, C. C., & Reddy, C. K. (2013). Data Clustering Algorithms and Applications. New York: Chapman and Hall.
- [3] Anon, n. (2021, January 12). Clustering Scikit Learn 0.24.0 Documentation. Retrieved from Scikit Learn: <https://scikit-learn.org/stable/modules/clustering.html#hierarchical-clustering>
- [4] Widarjono, A. (2010). Analisis Statistika Multivariat Terapan. Yogyakarta: UPP STIM YKPN.
- [5] Kurniawan, & Yuniarto. (2016). Analisis Regresi Dasar dan Penerapannya dengan R. Jakarta: Kencana.
- [6] Hanada, S. (2021). Penggunaan Analisis Cluster Dalam Pengelompokan Kecamatan di Kabupaten Karawang Berdasarkan Metode Kontrasepsi Peserta KB Aktif. Prosiding Statistika, 7(1), 42-49.
- [7] Shofwani Sheila Ghazia, Kudus Abdul. (2021). *Penentuan Kriteria Pengunjung dalam Pemilihan Green Hotel di Kota Bandung Menggunakan Metode Discrete Choice Experiment dengan Desain Choice Sets Kombinatorial*. Jurnal Riset Statistika, 1(1), 1-9.