

## PENGELOMPOKKAN KABUPATEN/KOTA DIJAWA TENGAH BERDASARKAN VARIABELINDIKATOR KESEHATAN MENGGUNAKAN ANALISIS CLUSTER

Riki Dwi Prasetyo<sup>1</sup>, Kris Suryowati<sup>2</sup>, Rokhana Dwi Bekti<sup>3</sup>

<sup>1,2,3</sup> Jurusan Statistika, FST, Institut Sains & Teknologi AKPRIND Yogyakarta  
e-mail : rikdwi@gmail.com

**Abstract.** *Health development is an effort to increase awareness, willingness and ability of healthy life for everyone to realize the degree of public health as high. Therefore information about the characteristics of health indicators in Central Java is important to determine the success of health development in Central Java Province. This research conduct an analysis to determine the level of health in Central Java Province, the factors that influence the level of health, with the grouping of districts / cities in Central Java by a variable indicator. The method used is descriptive analysis, factor analysis and cluster analysis. Factor analysis produce three factors: environmental and health facilities factor, health complaints factors and health of mothers and infants factors. Grouping using hierarchical cluster analysis, namely: complete linkage, average linkage and Ward. Of the three methods used to obtained the best method looked at the value of Wilk's Lambda MANOVA is the Ward Method. This method produces a three cluster, which consists of 16 districts / cities for cluster 1 with health quality is quite good, 4 districts / cities to cluster 2 with health quality is poor, and 15 districts / cities for cluster 3 with the quality of good health.*

**Keywords :** *health indicator, factor analysis, cluster analysis*

**Abstrak.** Pembangunan kesehatan adalah upaya untuk meningkatkan kesadaran, kemauan, dan kemampuan hidup sehat bagi setiap orang agar terwujud derajat kesehatan masyarakat yang setinggi-tingginya. Oleh karena itu, informasi tentang karakteristik indikator kesehatan di Jawa Tengah adalah hal penting untuk mengetahui tingkat keberhasilan pembangunan kesehatan khususnya di Provinsi Jawa Tengah. Penelitian ini melakukan analisis untuk mengetahui tingkat kesehatan di Provinsi Jawa Tengah, faktor yang berpengaruh pada tingkat kesehatan, serta pengelompokan kabupaten/kota di Jawa Tengah berdasarkan variabel indikatornya. Metode yang digunakan adalah analisis deskriptif, analisis faktor dan analisis kluster. Analisis faktor menghasilkan tiga faktor yaitu faktor lingkungan dan fasilitas kesehatan, faktor keluhan kesehatan dan faktor kesehatan ibu dan balita. Pengelompokan menggunakan analisis kluster hirarki, yaitu : *complete linkage*, *average linkage* dan *ward*. Dari ke tiga metode yang digunakan diperoleh metode yang terbaik dengan melihat nilai *Wilk's Lambda MANOVA* adalah metode *Ward*. Metode ini menghasilkan 3 kluster, yang terdiri dari 16 kabupaten/kota untuk kluster 1 dengan kualitas kesehatan cukup baik, 4 kabupaten/kota untuk kluster 2 dengan kualitas kesehatan kurang baik, dan 15 kabupaten/kota untuk kluster 3 dengan kualitas kesehatan baik.

**Kata kunci :** *indikator kesehatan, analisis faktor, analisis kluster*

### 1. Pendahuluan

Pembangunan kesehatan adalah upaya untuk meningkatkan kesadaran, kemauan, dan kemampuan hidup sehat bagi setiap orang agar terwujud derajat kesehatan masyarakat yang setinggi-tingginya. Melalui pembangunan kesehatan diharapkan akan mempunyai sumber daya manusia (SDM) yang berkualitas. Salah satu kualitas SDM yang baik adalah sehat jasmani dan rohani [1]. Pembangunan kesehatan di Jawa Tengah belum dapat dikatakan berhasil, karena masih tingginya angka kematian bayi dan angka kematian ibu. Padahal Jawa Tengah termasuk enam provinsi besar dengan pelayanan kesehatan yang baik. Angka kematian bayi dan ibu merupakan indikator utama kesehatan selain angka kesakitan dan status gizi[2].

Untuk menunjang keberhasilan pelaksanaan program pembangunan terutama yang berkaitan dengan kesehatan di Provinsi Jawa Tengah dan berdasarkan beberapa gambaran tentang kesehatan terdapat beberapa faktor yang mempengaruhinya berupa indikator kesehatan. Selanjutnya akan dilakukan analisis kluster untuk mengelompokkan kabupaten/kota di Provinsi Jawa Tengah berdasarkan variabel indikator kesehatan.

Penelitian sebelumnya yang membahas analisis faktor dan kluster tentang kesehatan diantaranya Jannah (2014), yang berjudul Analisis Pengelompokan Kecamatan di Surabaya Berdasarkan Indikator Pelayanan Kesehatan. Dalam hal ini metode yang digunakan adalah analisis faktor, analisis *cluster* dan analisis diskriminan.[3] Amaliyah (2012) yang berjudul Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Indonesia Sehat 2010, penelitian ini menggunakan metode *cluster* meliputi *single linkage*, *average linkage* dan *complete linkage*. [4]

Berdasarkan uraian di atas, maka pembahasan pada penelitian ini yaitu tentang deskripsi indikator-indikator kesehatan dan variabel apa saja yang berpengaruh terhadap kesehatan pada kabupaten/kota di Provinsi Jawa Tengah serta pengelompokan kabupaten/kota di Provinsi Jawa Tengah berdasarkan variabel-variabel yang mempengaruhinya.

## 2. Metode Penelitian

### Analisis Komponen Utama

Analisis komponen utama merupakan analisis statistika multivariat yang dapat digunakan untuk mereduksi sejumlah variabel asal menjadi beberapa variabel baru yang bersifat ortogonal dan tetap mempertahankan total keragaman dari peubah asalnya. Banyaknya komponen utama yang dipilih dapat ditentukan dengan beberapa prosedur yaitu berdasarkan *eigenvalues*, *scree plot* dan persentase keberagaman [5].

### Analisis Faktor

Analisis faktor adalah analisis statistika yang digunakan untuk mereduksi atau meringkas data dengan cara menganalisis sejumlah observasi dipandang dari interkorelasi atau keterkaitan antar variabel asli. Hal yang harus dipenuhi sebelum melakukan analisis faktor adalah uji kelayakan variabel dan adanya korelasi antar variabel. Uji yang dipakai untuk menilai layak tidaknya sebuah variabel untuk analisis faktor yaitu :

#### 1) Uji Bartlett

Tujuan uji *Bartlett* yaitu untuk menguji apakah matriks korelasi yang dihasilkan adalah matriks identitas. Matriks identitas mengindikasikan bahwa diantara variabel tidak terdapat korelasi. Statistik uji yang digunakan :

$$\chi^2_{hitung} = - \left[ (N - 1) - \frac{(2p+5)}{6} \right] \ln |R| \tag{1}$$

dengan N : jumlah observasi

|R| : determinan matriks korelasi

p : banyaknya variabel

#### 2) Uji Kaiser-Meyer-Olkin (KMO)

Uji *KMO* digunakan untuk mengetahui apakah data yang digunakan dapat dianalisis lebih lanjut atau tidak dengan analisis faktor. Statistik uji *KMO* adalah:

$$KMO = \frac{\sum_i \sum_{i \neq j} r_{ij}^2}{\sum_i \sum_{i \neq j} r_{ij}^2 + \sum_i \sum_{i \neq j} a_{ij}^2} \tag{2}$$

i : 1,2,...,p dan j : 1,2,...,p

dengan  $r_{ij}$  : koefisien korelasi antara peubah i dan j

$a_{ij}$  : koefisien korelasi parsial antara peubah i dan j

p : banyaknya variabel [5]

Ujilain yang digunakan untuk melihat korelasi antar variabel dan dapat tidaknya analisis faktor, dilakukan uji *Measure of Sampling Adequacy* (MSA). Apabila ada lebih dari satu variabel yang mempunyai nilai MSA dibawah 0,5 maka yang dikeluarkan adalah variabel dengan MSA terkecil. Kemudian proses pengujian kelayakan variabel diulang sampai didapat variabel yang layak dianalisis faktor sesuai kriteria uji. [6]

### Analisis Kluster

Analisis kluster adalah suatu metode statistik yang mengidentifikasi kelompok sampel berdasarkan karakteristik yang sama dengan tingkat homogenitas yang tinggi antar obyek menjadi kluster yang berbeda dengan tingkat heterogenitas obyek yang tinggi antar kluster.[5]Langkah-langkah analisis kluster meliputi :

- 1) Pemilihan Metode Pengukuran Jarak  
Kemiripan di antara dua objek ditunjukkan oleh jarak antar kedua objek tersebut. Pengukuran jarak yang digunakan yaitu jarak *euclidean* dan *squared euclidean*.
- 2) Pemilihan Prosedur Pengklasteran  
Ditinjau dari metode pengelompokannya, analisis kluster memiliki dua metode, yaitu metode hirarki dan non hirarki. Metode yang digunakan dalam penelitian ini adalah metode hirarki. Metode-metode pengelompokan hirarki dibedakan berdasarkan konsep jarak antar kelompok, penentuan jarak antar kelompok.
- 3) Menentukan Jumlah Kluster  
Pada kluster hirarki, jumlah kluster dapat ditentukan dengan melihat output *agglomeration* dan *dendrogram*.
- 4) Validasi Kluster  
Kluster yang telah terbentuk kemudian diuji apakah metode pengklasteran tersebut valid atau tidak. Metode validasi kluster yang digunakan yaitu *Multivariate Analysis of Variance* (MANOVA). Model MANOVA menggunakan variabel dalam analisis kluster sebagai variabel dependen dan keanggotaan kluster sebagai variabel independen. MANOVA dipilih karena variabel dependen saling berkorelasi [7]. Statistik uji yang digunakan yaitu nilai *Wilk's Lambda*. Tolak  $H_0$  jika nilai *Wilk's Lambda* lebih besar dari  $F_{(\alpha, df_1, df_2)}$ . [8]
- 5) Menginterpretasikan Kluster yang Terbentuk  
Interpretasi profil kluster meliputi pengkajian mengenai *centroid*. *Centroid* yaitu rata-rata nilai obyek yang terdapat dalam kluster pada tiap variabel. Nilai tersebut memungkinkan untuk menguraikan setiap kluster dengan cara memberi suatu label atau nama. [5]

### 3. Analisis dan Pembahasan

#### Analisis Deskriptif

Analisis dapat dilihat pada Tabel 1 dibawah ini :

Tabel 1. Statistik Deskriptif Indikator Kesehatan

No	Variabel	Min	Max	Mean	Standar Deviasi
1	Penduduk yang mengalami keluhan kesehatan ( $X_1$ )	20.65	56.12	32.13	8.56
2	Penduduk yang mengalami keluhan kesehatan dan berobat sendiri ( $X_2$ )	43.03	71.10	54.60	6.70
3	Penduduk yang mengalami keluhan kesehatan dan berobat jalan ( $X_3$ )	36.93	66.05	52.14	5.93
4	Penduduk yang berobat sendiri dan menggunakan obat tradisional ( $X_4$ )	3.75	13.68	7.09	2.56
5	Bayi lahir ditolong tenaga kesehatan ( $X_5$ )	70.62	100.00	94.09	6.69
6	Balita yang sudah imunisasi lengkap ( $X_6$ )	60.34	85.33	72.98	7.04
7	Balita yang diberi ASI tanpa makanan tambahan selama 0-4 bulan ( $X_7$ )	29.20	72.31	54.80	11.63
8	Rumah tangga yang menggunakan sumber air bersih ( $X_8$ )	46.39	98.43	70.94	11.75
9	Rumah tangga yang memiliki akses sanitasi layak ( $X_9$ )	15.57	92.61	68.35	17.45
10	Rumah tangga yang memiliki akses air minum layak ( $X_{10}$ )	55.58	98.84	72.98	11.17
11	Angka kematian bayi ( $X_{11}$ )	0.38	1.78	1.04	0.30

12 Angka kematian ibu ( $X_{12}$ ) 0.01 0.24 0.12 0.05

Tabel 1 menunjukkan penyebaran data untuk semua variabel relatif kecil. Penyebaran data tertinggi pada variabel rumah tangga yang memiliki akses sanitasi layak, karena memiliki *range* yang terbesar (77,04). Kualitas kesehatan bisa dikatakan baik apabila jumlah keluhan kesehatan, angka kematian bayi dan ibu yang seminimal mungkin. Sedangkan kesadaran pengobatan medis, penolong kelahiran, gizi balita dan fasilitas rumah tangga harus maksimal.

**Analisis Faktor**

Analisis faktor digunakan untuk mereduksi variabel yang bertujuan untuk mengatasi korelasi antar variabel yang dapat mengganggu proses pembentukan kelompok [1]. Tahappertamasebelummelakukananalisisfaktoryaitumenentukanvariabel yang layakdancukupuntkdianalisisfaktor.

Tabel 2. Uji Kelayakan Variabel

<i>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</i>	0,492
<i>Bartlett's Test of Sphericity</i>	<i>Approx. Chi-Square</i> 197,667
	<i>df</i> 66
	<i>Sig.</i> 0,00

Dari Tabel 2 didapatkan *p-value* kurang dari  $\alpha=0,05$  sehingga dapat disimpulkan antar variabel saling berkorelasi. Kemudian diperoleh nilai KMO 0,492 yang lebih kecil dari 0,5 sehingga dapat dikatakan bahwa jumlah data dari variabel-variabel indikator kesehatan belum memenuhi uji kecukupan data. Oleh karena itu, selanjutnya melihat nilai MSA masing-masing variabel pada Tabel 3.

Tabel 3. Nilai MSA Masing-Masing Variabel

Variabel	Nilai MSA
Penduduk yang mengalami keluhan kesehatan ( $X_1$ )	0,462
Penduduk yang mengalami keluhan kesehatan dan berobat sendiri ( $X_2$ )	0,493
Penduduk yang mengalami keluhan kesehatan dan berobat jalan ( $X_3$ )	0,436
Penduduk yang berobat sendiri dan menggunakan obat tradisional ( $X_4$ )	0,515
Bayi lahir ditolong tenaga kesehatan ( $X_5$ )	0,685
Balita yang sudah imunisasi lengkap ( $X_6$ )	0,450
Balita yang diberi ASI tanpa makanan tambahan selama 0-4 bulan ( $X_7$ )	0,315
Rumah tangga yang menggunakan sumber air bersih ( $X_8$ )	0,474
Rumah tangga yang memiliki akses sanitasi layak ( $X_9$ )	0,660
Rumah tangga yang memiliki akses air minum layak ( $X_{10}$ )	0,499
Angka kematian bayi ( $X_{11}$ )	0,229
Angka kematian ibu ( $X_{12}$ )	0,622

Tabel 3 menunjukkan beberapa variabel memiliki nilai MSA kurang dari 0,5, artinya variabel tidak bisa diprediksi dan tidak dapat dianalisis lebih lanjut. Karena terdapat lebih dari satu variabel yang memiliki nilai MSA kurang dari 0,5, maka selanjutnya memilih variabel dengan nilai MSA terkecil untuk dikeluarkan dari analisis yaitu variabel angka kematian bayi.

Setelah variabel angka kematian bayi dikeluarkan, maka dilakukan uji lagi dengan tujuan mendapatkan variabel yang layak untuk dianalisis selanjutnya dengan kriteria nilai KMO yang lebih dari 0,5 dan nilai MSA masing-masing variabel lebih dari 0,5. Hasil akhir nilai MSA masing-masing variabel yang ditunjukkan pada Tabel 4.

Tabel 4 Nilai Akhir Msa Masing-Masing Variabel

Variabel	Nilai MSA
Penduduk yang mengalami keluhan kesehatan ( $X_1$ )	0,610
Penduduk yang mengalami keluhan kesehatan dan berobat sendiri ( $X_2$ )	0,511
Penduduk yang mengalami keluhan kesehatan dan berobat jalan ( $X_3$ )	0,501

Bayi lahir ditolong tenaga kesehatan ( $X_5$ )	0,768
Balita yang sudah imunisasi lengkap ( $X_6$ )	0,501
Rumah tangga yang menggunakan sumber air bersih ( $X_8$ )	0,592
Rumah tangga yang memiliki akses sanitasi layak ( $X_9$ )	0,621
Rumah tangga yang memiliki akses air minum layak ( $X_{10}$ )	0,602
Angka kematian ibu ( $X_{12}$ )	0,690

Setelah semua variabel telah layak atau cukup untuk dianalisis faktor maka analisis selanjutnya yaitu menentukan banyaknya faktor yang terbentuk dengan melihat *eigenvalues*.

Tabel 5. Faktor yang Terbentuk

Faktorke-i	Akar Ciri/ <i>eigen value</i> ( $\lambda$ )	Persentase Keragaman (%)	Kumulatif (%)
1	2,950	32,777	32,777
2	2,117	23,527	56,304
3	1,438	15,979	72,282

Tabel 5 menunjukkan dari 9 variabel asli didapatkan 3 faktor dengan total varians kumulatif yaitu 72,282 persen. Untuk mengetahui variabel apa saja yang termasuk dalam faktor maka dapat dilihat dari nilai loading factor pada Tabel 6.

Tabel 6 Nilai *Loading Factor*

Variabel	Faktor 1	Faktor 2	Faktor 3
Penduduk yang mengalami keluhan kesehatan ( $X_1$ )	-0,044	<b>0,748</b>	0,186
Penduduk yang mengalami keluhan kesehatan dan berobat sendiri ( $X_2$ )	-0,153	<b>0,779</b>	-0,144
Penduduk yang mengalami keluhan kesehatan dan berobat jalan ( $X_3$ )	0,222	<b>-0,694</b>	<b>0,569</b>
Bayi lahir ditolong tenaga kesehatan ( $X_5$ )	<b>0,811</b>	-0,121	-0,054
Balita yang sudah imunisasi lengkap ( $X_6$ )	-0,056	-0,407	<b>-0,763</b>
Rumah tangga yang menggunakan sumber air bersih ( $X_8$ )	<b>0,871</b>	0,361	-0,050
Rumah tangga yang memiliki akses sanitasi layak ( $X_9$ )	<b>0,679</b>	-0,150	0,322
Rumah tangga yang memiliki akses air minum layak ( $X_{10}$ )	<b>0,869</b>	0,309	-0,021
Angka kematian ibu ( $X_{12}$ )	0,489	-0,204	<b>0,605</b>

Tabel 6 menunjukkan bahwa nilai *loading factor* untuk menentukan keanggotaan suatu faktor. Dari tabel tersebut didapat anggotafaktor 1 adalah  $X_5$ ,  $X_8$ ,  $X_9$  dan  $X_{10}$ . Variabel anggotafaktor 2 adalah  $X_1$  dan  $X_2$ . Anggota faktor 3 adalah  $X_6$  dan  $X_{12}$ . Variabel  $X_3$  memiliki korelasi yang tinggi pada 2 faktor yang berbeda yaitu faktor 2 dan 3. Masuknya variabel  $X_3$  pada 2 faktor menyebabkan hasil analisis menjadi tidak tepat. Oleh karena itu dilakukan proses rotasi faktor. Hasil rotasi faktor ditunjukkan oleh Tabel 7.

Tabel 7 Nilai *Loading Factor*  $X_3$  Setelah Dirotasi

Variabel	Faktor 1	Faktor 2	Faktor 3
$X_3$	0,099	<b>-0,892</b>	0,222

Tabel 7 menunjukkan nilai *loading factor*  $X_3$  setelah dirotasi menggunakan metode *varimax*. Variabel  $X_3$  yang sebelumnya berkorelasi kuat terhadap faktor 2 dan 3, setelah dirotasi menjadi lebih jelas masuk ke salah satu faktor, yaitu faktor 2.

Setelah diperoleh sejumlah faktor yang terbentuk dan anggotanya, langkah terakhir adalah interpretasi hasil analisis faktor yaitu penamaan pada ketiga faktor yang terbentuk.

- Faktor 1 terdiri dari bayi lahir ditolong tenaga kesehatan, rumah tangga yang menggunakan sumber air bersih, memiliki akses sanitasi layak dan air minum layak. Faktor ini diberi nama FAKTOR LINGKUNGAN DAN FASILITAS KESEHATAN.

- b) Faktor 2 terdiri dari penduduk yang mengalami keluhan kesehatan satu bulan terakhir, penduduk yang mengalami keluhan kesehatan dan berobat sendiri, dan penduduk yang mengalami keluhan kesehatan dan berobat jalan. Faktor ini diberi nama FAKTOR KELUHAN KESEHATAN
- c) Faktor 3 terdiri dari balita yang sudah imunisasi lengkap dan angka kematian ibu

**Analisis Kluster**

**Metode Complete Linkage**

Metode pengelompokkan dengan pendekatan jarak jauh. Dalam metode ini seluruh objek dalam suatu kluster dikaitkan satu sama lain pada suatu jarak maksimum atau dengan kesamaan minimum. Berikut hasil pengelompokkan dengan metode *complete linkage* :

- 1) Mengukur Kesamaan Objek.

Kesamaan objek diukur dari jarak *euclidean*. Sebagai contoh jarak antara Cilacap dengan Banyumas adalah 28,745, sedangkan jarak Cilacap dengan Kebumen adalah 31,124. Hal ini berarti Cilacap lebih mirip dalam karakteristik kesehatannya dengan Banyumas dibandingkan dengan Kebumen. Perhitungan jarak euclidean antara Kabupaten Cilacap dan Kabupaten Banyumas adalah :

$$d_{ab} = \sqrt{\sum_{k=1}^p (x_{ak} - x_{bk})^2}$$

$$d_{12} = \sqrt{(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 + \dots + (x_{19} - x_{29})^2}$$

$$= \sqrt{(31,08 - 31,50)^2 + (57,69 - 64,96)^2 \dots + (0,12 - 0,115)^2}$$

$$= 28,745$$

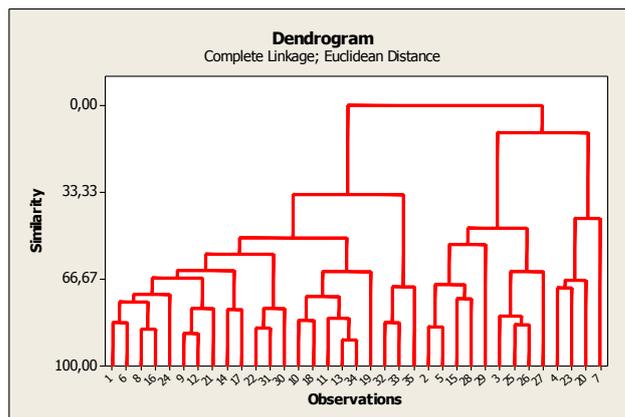
Jadi jarak antara Cilacap dan Banyumas adalah sebesar 28,745.

- 2) Membentuk Kluster

Setelah jarak antar variabel diukur dengan jarak euclidean, maka dilakukan pengelompokan, yang dilakukan secara bertingkat. Penggabungan anggota-anggota kluster dapat dilihat dari *Agglomeration Schedule*.

- 3) Menentukan Banyaknya Kluster

Dendrogram menunjukkan anggota kluster dalam bentuk diagram batang atau diagram pohon. Dendrogram untuk metode *Complete Linkage* ditunjukkan oleh Gambar 1.



Gambar 1 Dendrogram Metode *Complete Linkage*

Dendrogram menunjukkan gambaran proses kluster yang terjadi. Banyaknya kluster ditentukan yaitu 3, 4 dan 5 kluster. Untuk menguji banyaknya kluster yang paling tepat digunakan uji MANOVA. Hasil uji MANOVA ditunjukkan oleh Tabel 4.11 sebagai berikut :

Tabel 8 Penentuan Jumlah Kluster Metode *Complete Linkage*

Jumlah Kluster Ter rbentuk	Uji MANOVA			Kesimpulan
	Nilai Wilk's Lambda	Nilai F <sub>tabel</sub>	p-value	
3	7,392	1,82	0,000	Ada perbedaan rata-rata antarkluster yang terbentuk.
4	5,806	1,65	0,000	
5	4,731	1,56	0,000	

Tabel 8 menunjukkan jumlah kluster yang paling tepat untuk mengelompokkan data dengan metode *complete linkage*. Semakin besar nilai *Wilk's Lambda* maka banyaknya kluster yang terbentuk semakin bagus. Oleh karena itu, dapat disimpulkan bahwa 3 kluster merupakan jumlah terbaik untuk metode *complete linkage*. Masing-masing anggota kluster 1, 2 dan 3 sebagai berikut :

1. Anggota kluster 1 meliputi Kabupaten Cilacap, Purworejo, Magelang, Boyolali, Klaten, Sukoharjo, Wonogiri, Karanganyar, Sragen, Blora, Rembang, Pati, Kudus, Demak, Semarang, Kendal, Kota Magelang, Surakarta, Semarang, Salatiga, Pekalongan dan Tegal.
2. Anggota kluster 2 meliputi Kabupaten Batang, Pekalongan, Pemalang, Tegal, Brebes, Banyumas, Purbalingga, Kebumen dan Grobogan
3. Anggota Kluster 3 meliputi Kabupaten Banjarnegara, Wonosobo, Temanggung dan Jepara.

### Metode Average Linkage

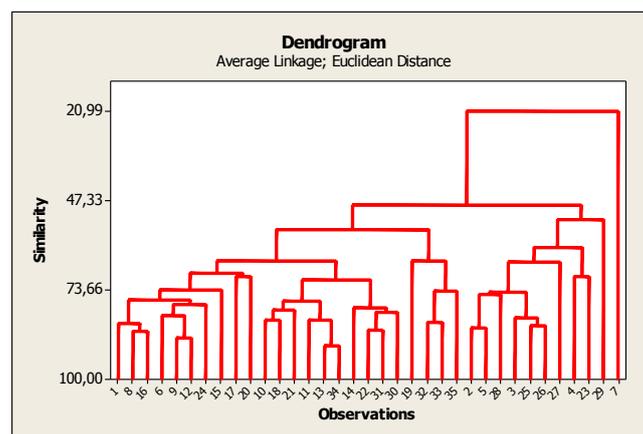
Jarak antara dua kluster pada metode *average linkage* didefinisikan sebagai rata-rata jarak antara semua pasangan objek dengan salah satu pasangan berasal dari setiap kluster. Berikut hasil pengelompokkan dengan menggunakan metode *average linkage* :

#### 1) Membentuk Kluster

Setelah jarak antar variabel diukur dengan jarak *euclidean*, maka dilakukan pengelompokkan, yang dilakukan secara bertingkat. Penggabungan anggota-anggota kluster dapat dilihat dari *Agglomeration Schedule*.

#### 2) Menentukan Banyaknya Kluster

Alat yang membantu untuk menentukan jumlah kluster disebut *dendrogram*. *Dendrogram* menunjukkan anggota kluster dalam bentuk diagram batang atau diagram pohon. *Dendrogram* untuk metode *Average Linkage* ditunjukkan oleh Gambar 2.



Gambar 2 Dendrogram Metode *Average Linkage*

*Dendrogram* menunjukkan gambaran proses kluster yang terjadi. Banyaknya kluster ditentukan yaitu 3, 4 dan 5 kluster. Untuk menguji banyaknya kluster yang paling tepat digunakan uji MANOVA. Hasil uji MANOVA ditunjukkan oleh Tabel 9.

Tabel 9 Penentuan Jumlah Kluster Metode *Average Linkage*

JumlahKlusterTerbentuk	UjiMANOVA			Kesimpulan
	NilaiWilk's Lambda	NilaiF <sub>tabel</sub>	P- value	
3	5,183	1,82	0,000	Ada perbedaan rata-rata anatar kluster yang terbentuk.
4	3,840	1,65	0,000	
5	4,926	1,56	0,000	

Tabel 9 menunjukkan bahwa jumlah kluster yang paling tepat untuk mengelompokkan data dengan metode *average linkage*. Oleh karena itu, dapat disimpulkan bahwa 3 kluster merupakan jumlah terbaik untuk metode *average linkage*.

Masing-masing anggota kluster 1, 2 dan 3 sebagai berikut :

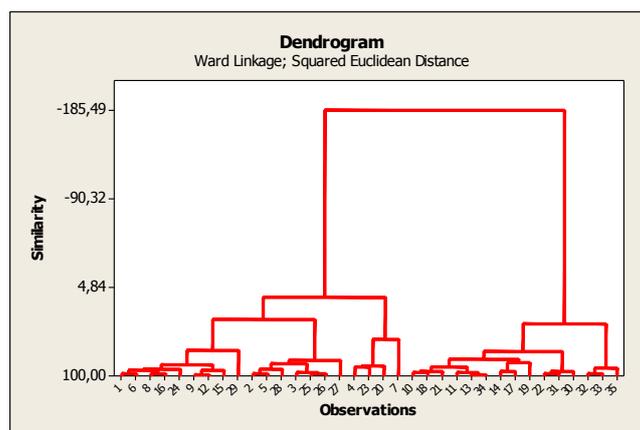
1. Anggotakluster 1 meliputi Kabupaten Cilacap, Purworejo, Magelang, Boyolali, Klaten, Sukoharjo, Wonogiri, Karanganyar, Sragen, Grobogan, Blora, Rembang, Pati, Kudus, Jepara, Demak, Semarang, Kendal, Kota Magelang, Surakarta, Semarang, Salatiga, Pekalongan dan Tegal.
2. Anggotakluster 2 meliputi Kabupaten Banyumas, Purbalingga, Banjarnegara, Kebumen, Temanggung, Batang, Pekalongan, Pemalang, Tegal dan Brebes.
3. Anggota Kluster 3 meliputi Kabupaten Wonosobo.

**Metode Ward**

Metode *ward* didasarkan pada *sum square error* (SSE) dengan ukuran kehomogenan antara dua objek berdasarkan jumlah kuadrat kesalahan yang paling minimal. Berikut hasil pengelompokkan dengan metode *ward* :

- 1) Mengukur Kesamaan Objek.  
Kesamaan objek diukur dari jarak *squared Euclidean*.
- 2) Membentuk Kluster  
Setelah jarak antar variabel diukur dengan jarak *squared euclidean*, maka dilakukan pengelompokkan, yang dilakukan secara bertingkat. Penggabungan anggota-anggota kluster dapat dilihat dari *Agglomeration Schedule*.
- 3) Menentukan Banyaknya Kluster

Kluster yang terbentuk dapat dilihat pada gambar 3. Hasil yang didapat dari dendrogram, tampak bahwa data yang dianalisis akan tergabung menjadi satu *cluster*. Berikut *dendrogram* hasil pengelompokkan dengan metode *ward* :



Gambar 3 Dendrogram Metode *Ward*

Dendrogram menunjukkan gambaran proses kluster yang terjadi. Banyaknya kluster ditentukan yaitu 3, 4 dan 5 kluster. Untuk menguji banyaknya kluster yang paling tepat digunakan uji MANOVA. Hasil uji MANOVA ditunjukkan oleh Tabel 10.

Tabel 10 Penentuan Jumlah Kluster Metode Ward

Jumlah Kluster Terbentuk	Uji MANOVA			Kesimpulan
	Nilai Wilk's Lambda	Nilai F <sub>tabel</sub>	p-value	
3	9,019	1,82	0,000	Ada perbedaan rata-rata antar kluster yang terbentuk.
4	6,379	1,65	0,000	
5	5,871	1,56	0,000	

Dari Tabel 10 dapat disimpulkan bahwa 3 kluster merupakan jumlah terbaik untuk metode *complete linkage*. Masing-masing anggota kluster 1, 2 dan 3 sebagai berikut :

1. Anggota kluster 1 meliputi Kabupaten Cilacap, Banyumas, Purbalingga, Kebumen, Purworejo, Magelang, Boyolali, Wonogiri, Grobogan, Blora, Kendal, Batang, Pekalongan, Pemalang, Tegal dan Brebes.
2. Anggota kluster 2 meliputi Kabupaten Banjarnegara, Wonosobo, Jepara dan Temanggung.
3. Anggota kluster 3 meliputi Kabupaten Klaten, Sukoharjo, Karanganyar, Sragen, Rembang, Pati, Kudus, Demak, Semarang, Kota Magelang, Surakarta, Salatiga, Semarang Pekalongan dan Tegal.

### Penentuan Metode Kluster Terbaik

Metode penentuan kluster terbaik yang digunakan yaitu MANOVA. Hasil uji MANOVA ditunjukkan oleh Tabel 11 sebagai berikut :

Tabel 11 Penentuan Kluster Terbaik Menggunakan Metode MANOVA

Metode Pengklasteran	Jumlah Kluster Terbentuk	Uji MANOVA		
		Nilai Wilk's Lambda	Nilai F <sub>tabel</sub>	p-value
<i>Complete Linkage</i>	3	7,293	1,82	0,000
<i>Average Linkage</i>	3	5,183	1,82	0,000
<i>Ward</i>	3	9,019	1,82	0,000

Pada Tabel 11 di atas dapat diketahui bahwa nilai *Wilk's Lambda* metode Ward lebih besar dibandingkan dengan metode *complete* dan *average linkage* sehingga metode yang lebih baik dan tepat digunakan dalam penelitian ini yaitu metode *Ward*.

### Interpretasi dan Profiling Kluster

Pada tahap interpretasi meliputi pengujian pada masing-masing kluster yang terbentuk untuk memberikan nama atau keterangan secara tepat sebagai gambaran sifat dari kluster tersebut. Ketika memulai proses interpretasi digunakan rata-rata (*centroid*) setiap kluster pada setiap variable. Hasil rata-rata setiap variabel yang berpengaruh dari metode *Ward* ditunjukkan Tabel 12.

Tabel 12 Interpretasi dan Pembuatan Profil

Kluster	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	X <sub>12</sub>
1	33,22	55,86	50,54	90,28	73,75	63,20	62,75	65,03	0,130
2	35,24	53,95	47,62	92,56	74,43	71,35	34,39	73,26	0,106
3	30,13	53,44	55,05	98,56	71,78	79,09	83,38	81,38	0,109

Berdasarkan Tabel 12 kualitas kesehatan tertinggi pertama (merah) berada pada kluster 3. Kualitas kesehatan tertinggi kedua (hitam) berada pada kluster 2 dan kualitas kesehatan tertinggi ketiga (hijau) berada pada kluster 1.

Berdasarkan hasil pengelompokan yang telah dilakukan maka berikut gambar wilayah pengelompokan kabupaten/kota di Jawa Tengah berdasarkan indikator kesehatan dengan metode *Ward* :

