

Pengklasteran Kabupaten/Kota di Jawa Tengah berdasarkan Tenaga Kesehatan dengan Menggunakan Metode Ward dan K-Means

Sri Puji Lestari, Epha Diana Supandi, dan Pipit Pratiwi Rahayu

Program Studi Matematika Fakultas Sains dan Teknologi, UIN Sunan Kalijaga, Jl. Marsda Adisucipto No. 1 Yogyakarta, Indonesia

Korespondensi; Sri Puji Lestari, Email: sripujilestari511@gmail.com; Epha Diana Supandi, Email: epha.supandi@uin-suka.ac.id; Pipit Pratiwi Rahayu, Email: pipitrahayu@yahoo.com

Abstrak

Analisis kluster merupakan suatu metode yang digunakan untuk mengelompokkan objek (kasus) ke dalam kluster (kelompok) yang relatif sama. Tujuan penelitian ini untuk mengklasterkan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan tenaga kesehatan tahun 2015 seperti tenaga medis, tenaga keperawatan, tenaga kebidanan, tenaga kefarmasian dan tenaga kesehatan lainnya dengan menggunakan metode *Ward* dan *K-Means*. Hasil penelitian menunjukkan ada tiga kluster terbentuk dimana metode *Ward* menghasilkan nilai rasio simpangan baku sebesar 0,3019% lebih besar jika dibandingkan dengan nilai rasio simpangan baku pada metode *K-Means* yaitu 0,2974%. Pada kasus ini, metode *K-Means* merupakan metode yang lebih baik dibandingkan metode *Ward*.

Kata Kunci: Analisis kluster; metode *K-Means*; metode *Ward*; tenaga kesehatan

Abstract

Cluster analysis is a method used to group objects (cases) into clusters (groups) that are relatively the same. The purpose of this study is to classify districts/cities in Central Java Province based on health worker in 2015 such as medical personnel, nursing staff, midwifery staff, pharmacy personnel and health workers using the Ward and K-Means methods. The results show that there are three clusters formed where the Ward method produce a standard deviation ratio of 0.3019% greater than the standard deviation ratio in the K-Means method, which is 0.2974%. In this case, the K-Means method is a better method than the Ward method

Keywords: Cluster Analysis; K-Means Method; Ward Method; health worker

Pendahuluan

Analisis kluster merupakan suatu metode, yang dipergunakan untuk mengklasifikasi atau mengelompokkan objek (kasus) seperti orang, produk (barang), toko, perusahaan ke dalam kluster (kelompok). (Supranto, 2010).

Penerapan analisis kluster dapat ditemui dalam berbagai bidang antara lain bidang pemasaran contohnya pembentukan segmen berdasarkan data demografi, *psychographic profiles* mengenali *test market cities*, bidang asuransi contohnya mengelompokkan responden berdasarkan alasan dalam memutuskan untuk membeli suatu produk asuransi. Bidang pendidikan contohnya dalam mengelompokkan mahasiswa pemohon beasiswa, bidang kependudukan contohnya pengelompokkan desa tertinggal di suatu Kabupaten/Kota. Bidang pangan contohnya mengelompokkan potensi produksi buah-buahan di suatu Provinsi. Penerapan analisis kluster dapat dilihat Sitepu dkk (2011), Aftina dkk (2012) serta Yim dan Ramdeen (2015).

Penerapan analisis kluster pada bidang kesehatan diantaranya mengelompokkan suatu wilayah berdasarkan tingkat kesehatan. Tenaga kesehatan menurut Undang-Undang adalah setiap orang yang mengabdikan diri dalam bidang kesehatan serta memiliki pengetahuan dan/atau keterampilan melalui pendidikan di bidang kesehatan yang untuk jenis tertentu memerlukan kewenangan untuk melakukan

upaya kesehatan. Peran penting tenaga kesehatan itu sendiri dalam Undang-Undang Republik Indonesia nomor 36 tahun 2014 adalah untuk meningkatkan kualitas pelayanan kesehatan yang maksimal kepada masyarakat.

Tingkat penyebaran tenaga kesehatan di Indonesia belum merata ke seluruh wilayah terutama di Provinsi Jawa Tengah. Oleh karena itu, tujuan penelitian ini ingin mengelompokkan (mengklasterkan) Kabupaten/Kota di Jawa Tengah berdasarkan tenaga kesehatan seperti tenaga medis, tenaga keperawatan, tenaga kebidanan, tenaga kefarmasian dan tenaga kesehatan lainnya pada tahun 2015.

Landasan Teori

Analisis Klaster

Analisis klaster adalah teknik statistika yang digunakan untuk mengidentifikasi objek dengan memperhatikan beberapa kriteria. Analisis klaster termasuk analisis multivariat yang mewakili seluruh hubungan interdependensi, tidak ada perbedaan variabel bebas dan tak bebas dalam analisis ini (Supranto, 2010).

Prosedur pengklasteran dapat dilakukan dengan dua metode yaitu metode hirarki dan metode non-hirarki. Metode hirarki adalah teknik klastering yang membentuk tingkatan tertentu seperti struktur pohon dengan proses pengelompokkannya dilakukan secara bertingkat atau bertahap (Rafsanjadi dkk, 2012). Salah satu metode didalam kelas hirarki adalah metode *Ward*.

Metode *Ward*

Metode *Ward* (*Wards method*) adalah salah satu metode klaster dimana metode ini menggunakan *error sum of square (ESS)* sebagai pertimbangan dalam menggabungkan objek ke dalam klaster sehingga variansi di dalam klaster sekecil mungkin (Gudono, 2014).

Dimana rumus *ESS* adalah:

$$ESS = (\sum_{i=1}^n x_{iy}^2 - (\sum_{i=1}^n \bar{x}_{iy})^2) \quad (1)$$

Dimana,

x_{iy} : Nilai untuk objek ke- i pada klaster ke- y

n : Banyaknya objek dalam klaster yang terbentuk.

Misalkan diketahui sekumpulan data x_1, x_2, \dots, x_N dimana $x \in \mathfrak{R}^p$ maka langkah-langkah dalam metode *Ward* adalah (Johnson dan Wichern, 2001)

1. Setiap objek/pengamatan dianggap sebagai klaster, maka pada tahap ini mempunyai N klaster dengan *ESS* bernilai nol.
2. Selanjutnya terbentuk klaster dengan memilih dua dari N klaster yang memiliki nilai *ESS* terkecil. Secara sistematis, dari N klaster akan berkurang 1 pada setiap tahap ($N - 1$).
3. Membuat kombinasi dua pasang klaster baru yang terdiri dari satu klaster yang telah terbentuk dan klaster yang lain, kemudian menghitung kembali *ESS* dari setiap pasang klaster baru dan memilih dua pasang klaster yang memiliki nilai *ESS* terkecil untuk digabungkan menjadi satu klaster.
4. Mengulangi langkah ke-3 sampai semua objek bergabung menjadi satu klaster.

Metode *K-Means*

Menurut Johnson dan Wichern (2001) teknik pengelompokkan non-hirarki di rancang untuk mengelompokkan objek penelitian menjadi K klaster. Jumlah K klaster harus ditentukan terlebih dahulu oleh peneliti. Metode non-hirarki ini menandai setiap objek masuk ke dalam klaster yang mempunyai pusat klaster (rata-rata) terdekat. Metode non-hirarki sering disebut sebagai metode *K-Means*. Pada dasarnya metode *K-Means* berusaha untuk meminimalkan variasi dalam satu klaster dan memaksimalkan variasi antar klaster.

Secara umum analisis klaster metode *K-Means* menggunakan algoritma sebagai berikut:

1. Tentukan K sebagai jumlah klaster yang dibentuk.
2. Menentukan *centroid* (titik pusat klaster)

Penentuan *centroid* awal dilakukan secara random/acak dari objek-objek yang tersedia sebanyak K klaster, kemudian menghitung *centroid* klaster ke- i berikutnya, dengan rumus sebagai berikut:

$$C_{yk} = \frac{\sum_{k=1}^p x_k}{n} \quad (2)$$

Dimana:

C_{yk} : Nilai *centroid* dari klaster ke- y pada variabel ke- k

x_k : Nilai data pada variabel ke- k

n : Banyaknya objek dalam klaster yang terbentuk.

p : Banyaknya variabel yang diukur

- Hitung jarak setiap objek ke masing-masing *centroid* dari masing-masing klaster. Untuk menghitung jarak antar objek dengan *centroid* menggunakan jarak *Squared Euclidian*:

$$d_{(iy)} = \sum_{k=1}^p (x_{ik} - C_{yk})^2 \quad (3)$$

Dimana:

$d_{(iy)}$: Jarak *Squared Euclidean* antar objek ke- i dengan *centroid* klaster ke- y

p : Jumlah variabel klaster

x_{ik} : Nilai atau data dari objek ke- i pada variabel ke- k

C_{yk} : Nilai dari *centroid* klaster ke- y pada variabel ke- k .

- Alokasikan masing-masing objek kedalam *centroid* yang paling dekat.
- Lakukan iterasi kemudian tentukan posisi *centroid* baru dengan menggunakan persamaan (2).
- Ulangi langkah ke-3 sampai posisi *centroid* baru yang terbentuk nilainya sama (tidak berubah).

Bahan dan Metode

Objek-objek yang digunakan dalam penelitian ini adalah:

- | | | |
|----------------------|---------------------|---------------------|
| 1. Kab. Cilacap | 14. Kab. Sragen | 27. Kab. Pemalang |
| 2. Kab. Banyumas | 15. Kab. Grobogan | 28. Kab. Tegal |
| 3. Kab. Purbalinga | 16. Kab. Blora | 29. Kab. Brebes |
| 4. Kab. Banjarnegara | 17. Kab. Rembang | 30. Kota Magelang |
| 5. Kab. Kebumen | 18. Kab. Pati | 31. Kota Surakarta |
| 6. Kab. Purworejo | 19. Kab. Kudus | 32. Kota Salatiga |
| 7. Kab. Wonosobo | 20. Kab. Jepara | 33. Kota Semarang |
| 8. Kab. Magelang | 21. Kab. Demak | 34. Kota Pekalongan |
| 9. Kab. Boyolali | 22. Kab. Semarang | 35. Kota Tegal |
| 10. Kab. Klaten | 23. Kab. Temanggung | |
| 11. Kab. Sukoharjo | 24. Kab. Kendal | |
| 12. Kab. Wonogiri | 25. Kab. Batang | |
| 13. Kab. Karanganyar | 26. Kab. Pekalongan | |

Pada penelitian ini terdapat 5 tenaga kesehatan yang akan digunakan sebagai variabel dalam mengklasterkan Kabupaten/Kota di Jawa tengah tahun 2015 yaitu:

X_1 : Tenaga medis (*Medical personnel*)

X_2 : Tenaga keperawatan (*Nursing personnel*)

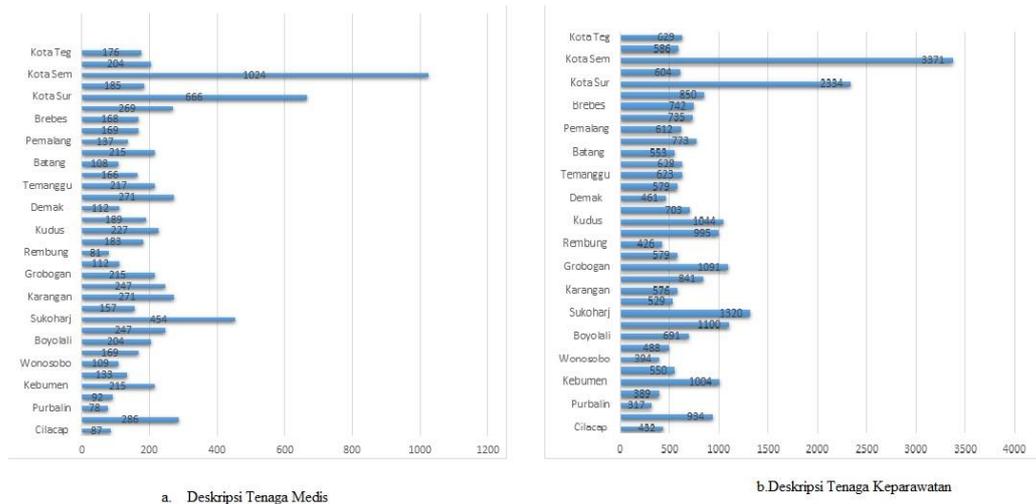
X_3 : Tenaga kebidanan (*Midwifery personnel*)

X_4 : Tenaga kefarmasian (*Pharmacy Personnel*)

X_5 : Tenaga kesehatan lainnya (*Other Health*)

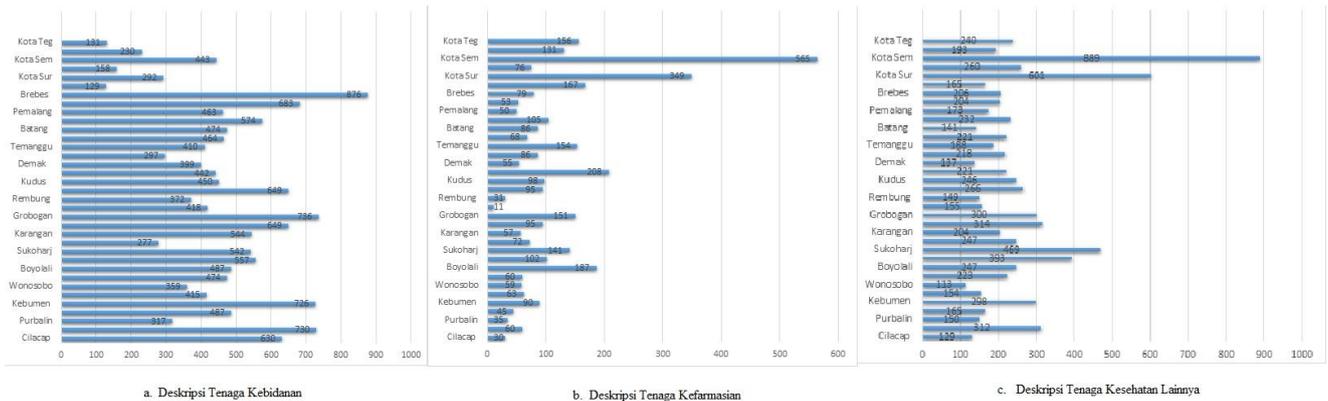
Hasil dan Pembahasan

Deskripsi jumlah tenaga kesehatan di Kabupaten/Kota Provinsi Jawa Tengah tahun 2015 untuk masing-masing variabel dapat dilihat pada gambar 1 dan 2 dibawah ini:



Gambar 1. Histogram Variabel Tenaga Medis dan Tenaga Keperawatan

Terlihat dalam gambar 1 pada panel a) dan b), jumlah tenaga medis dan tenaga keperawatan di kabupaten Semarang dan Surbaya memiliki nilai paling besar dibandingkan dengan kabupaten/kota lainnya.



Gambar 2. Histogram Variabel Tenaga Kesehatan, Tenaga Kefarmasian dan Tenaga Kesehatan Lainnya

Berdasarkan gambar 2 dapat dilihat bahwa jumlah tenaga kebidanan tertinggi terdapat pada Kabupaten Brebes, jumlah terbanyak untuk tenaga kefarmasian dan kesehatannya lainnya berada pada Kota Semarang yang juga merupakan ibukota provinsi Jawa Tengah.

Hasil kluster dengan menggunakan metode Ward terbentuk tiga klaser sebagai berikut:

1. Kluster pertama: terdiri dari 2 Kabupaten/Kota dimana pada kluster ini memiliki rata-rata tenaga kesehatan yang paling tinggi. Adapun anggota kluster pertama adalah:

Tabel 1. Kluster pertama dengan Menggunakan Metode Ward.

Kabupaten/Kota	Rata-rata Variabel
1. Kota Surakarta	848,4
2. Kota Semarang	1258,4
Rata-rata	1053,4

2. Kluster kedua: beranggotakan 11 Kabupaten/Kota dimana kluster kedua ini memiliki rata-rata tenaga kesehatan lebih rendah daripada kluster satu. Adapun anggota kluster kedua adalah:

Tabel 2. Kluster kedua dengan menggunakan Metode Ward.

Kabupaten/Kota	Rata-rata Variabel
1. Banyumas	464,4
2. Kebumen	466,6
3. Klaten	479,8
4. Sukoharjo	585,2
5. Sragen	429,2
6. Grobogan	498,6
7. Pati	437,6
8. Kudus	413
9. Pekalongan	379,8
10. Tegal	368,8
11. Brebes	414,2
Rata-rata	448,8363

3. Kluster tiga: kluster yang beranggotakan 22 Kabupaten/Kota dimana kluster ini memiliki rata-rata tenaga kesehatan paling rendah dibanding rata-rata variabel pada kluster pertama dan kedua. Anggota kluster ketiga adalah:

Tabel 3. Kluster ketiga dengan menggunakan Metode Ward.

Kabupaten/Kota	Rata-rata Variabel	Kabupaten/Kota	Rata-rata Variabel
1. Cilacap	261,6	2. Jepara	352,6
3. Purbalinga	179,4	4. Demak	232,8
5. Banjarnegara	235,6	6. Semarang	290,2
7. Purworejo	263	8. Temanggung	318,4
9. Wonosobo	206,8	10. Kendal	309,4
11. Magelang	282,8	12. Batang	272,4
13. Boyolali	363,2	14. Pemalang	287
15. Wonogiri	256,4	16. Kota Magelang	316
17. Karanganyar	330,4	18. Kota Salatiga	256,6
19. Bora	255	20. Kota Pekalongan	268,8
21. Rembung	211,8	22. Kota Tegal	266,4
Rata-rata		273,48	

Hasil pengklasteran dengan metode Ward dapat digambarkan pada gambar 4 berikut ini:



Gambar 4. Hasil pengklasteran dengan menggunakan Metode Ward.

Gambar 4 memperlihatkan kelompok Kabupaten/Kota dengan tingkat kualitas kesehatan yang tinggi hingga rendah berturut-turut adalah klaster 1, klaster 2, dan klaster 3, yang digambarkan pada peta Provinsi Jawa Tengah.

Hasil pengklasteran dengan menggunakan metode *K-Means* menghasilkan tiga klaster dengan rata-rata tiap variabel sebagai berikut:

Tabel 4. Rata-rata variabel pada tiap klaster dengan metode K-Means.

Variabel	Klaster 1	Klaster 2	Klaster 3
Tenaga Medis	186	666	1024
Tenaga Keperawatan	690	2334	3371
Tenaga Kebidanan	471	292	443
Tenaga Kefarmasian	90	349	565
Tenaga Kesehatan Lainnya	222	601	889

Hasil pengklasteran dapat digambarkan pada peta Provinsi Jawa Tengah berikut ini:



Gambar 5. Hasil pengklasteran dengan menggunakan Metode K-Means.

Berdasarkan gambar 5 dapat dilihat bahwa klaster pertama beranggotakan 33 Kabupaten/Kota, klaster dua beranggotakan 1 Kabupaten/Kota dan klaster tiga beranggotakan 1 Kabupaten/Kota. Anggota masing masing klaster adalah:

1. Klaster satu mempunyai anggota Cilacap, Banyumas, Purbalinga, Banjarnegara, Kebumen, Purworejo, Wonosobo, Magelang, Boyolali, Klaten, Sukoharjo, Wonogiri, Karanganyar, Sragen, Grobogan, Blora, Rembung, Pati, Kudus, Jepara, Demak, Semarang, Temanggung, Kendal, Batang, Pekalongan, Pemalang, Tegal, Brebes, Kota Magelang, Kota Salatiga, Kota Pekalongan, Kota Tegal.
2. Klaster dua beranggotakan Kota Surakarta.
3. Klaster tiga beranggotakan Kota Semarang.

Hasil pengklasteran dengan menggunakan metode *Ward* dan *K-Means* sama sama menghasilkan tiga klaster yaitu klaster yang mempunyai tingkat pelayanan tenaga kesehatan yang sangat baik (tinggi), sedang dan rendah.

Selanjutnya, untuk mengetahui kinerja kedua metode pengklasteran tersebut digunakan kriteria dua simpangan baku yaitu: simpangan baku dalam klaster (S_w) dan simpangan baku antar antar klaster (S_b). Suatu metode terbaik apabila mempunyai nilai rasio antara S_w dan S_b paling kecil artinya klaster tersebut mempunyai homogenitas yang tinggi.

Hasil perhitungan S_w dan S_b disajikan pada tabel berikut ini:

Tabel 5. Rasio antara Simpangan Baku Dalam dan Simpangan Baku Antar Klaster.

Simpangan Baku	Metode Ward	Metode K-Means
Simpangan baku dalam klaster (S_w)	156,0499	154,7042
Simpangan baku antar klaster (S_b)	516,8633	520,0794
Rasio $\frac{S_w}{S_b}$	0,3019 %	0,2974%

Pada tabel 5 terlihat bahwa simpangan baku dalam klaster (S_w) metode Ward lebih besar daripada metode K-Means. Sedangkan simpangan antar klaster (S_b) metode Ward lebih kecil daripada metode K-Means. Oleh karena itu, pada kasus ini metode K-Means lebih baik dalam mengklasterkan Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan variabel tenaga kesehatan dibandingkan metode Ward.

Kesimpulan

Hasil pengklasteran Kabupaten/Kota di Provinsi Jawa Tengah berdasarkan variabel tenaga kesehatan dengan menggunakan metode Ward dan K-Means menghasilkan tiga klaster. Pada kasus ini, pengklasteran dengan metode K-Means lebih baik dibandingkan metode Ward karena menghasilkan rasio antara S_w dan S_b paling kecil. Dengan kata lain, pengklasteran dengan metode K-Means menghasilkan tingkat homogenitas dalam klaster yang paling kecil (mirip) yang sesuai dengan tujuan analisis klaster.

Referensi

- [1] Gudono. 2014. Analisis Data Multivariat. Yogyakarta: BPFE.
- [2] J. Supranto. 2010. Analisis Multivariat Arti dan Interpretasi. Jakarta, Rineka Cipta.
- [3] R.A. Johnson and D.W. Wichern. 2001. Applied Multivariate Statistical Analysis (5th Edition). New York, Prentice Hall.
- [4] Rafsanjani, M Kuchaki., Varzaneh, Z Asghari., and Chukanlo, N Emami. (2012). A Survey of Hierarchical Clustering Algorithms. The Journal of Mathematics and Computer Science, 5(3), 229-240.
- [5] Sitepu, Robinson., Irmeilyana., dan Gultom, Berry. (2011). Analisis Cluster terhadap Tingkat Pencemaran Udara pada Sektor Industri di Sumatera Selatan. Jurnal Penelitian Sains, 14, 14303 11-14303 17.
- [6] T. Aftina, B. Santosa dan A.L Barakbah. (2012). Analisa Perbandingan Metode Hierarchical Clustering, K-means dan Gabungan Keduanya dalam Cluster Data (Studi Kasus: Problem Kerja Praktek Jurusan Teknik Industri ITS). Jurnal Teknik ITS. 1(1), A 521 A 525.
- [7] Yim, Odilia., & Ramdeen, Kylee T. (2015). Hierarchical Cluster Analysis: Comparison of Three Linkage Measures and Application to Psychological Data. The Quantitative Methods for Psychology, 11(1), 8-21.

