
Pengelompokkan Kabupaten/Kota Di Provinsi Jawa Timur Berdasarkan Indikator Kesehatan Mortalitas Dengan Algoritma Variable Weighting K-Means

Elvira Mustikawati Putri Hermanto
Universitas PGRI Adi Buana Surabaya
Jl. Dukuh Menanggal XII Telp./Fax. (031) 822289873 Surabaya 60234
e-mail: elvira.mustikawati@gmail.com

ABSTRAK

Salah satu indikator yang menggambarkan derajat kesehatan, pembangunan dan pelayanan kesehatan adalah angka kematian (mortalitas). Mortalitas dibentuk dari beberapa indikator, yaitu angka kelahiran kasar, angka kematian kasar, angka kematian bayi, angka kematian balita dan angka kematian ibu. Pada Tahun 2014, capaian angka kematian bayi masih berada diatas target MDGs yang telah disepakati. Distribusi indikator kesehatan, mortalitas, di Provinsi Jawa Timur perlu diketahui untuk mencapai target dalam MDGs dan untuk meningkatkan derajat kesehatan di Provinsi Jawa Timur. Penelitian ini membahas tentang pengelompokkan kabupaten-kabupaten (dan kota) yang ada di Provinsi Jawa Timur berdasarkan kemiripan lima indikator mortalitas dengan menggunakan metode Variable Weighted K-Means (VW K-Means). Berdasarkan metode silhouette diperoleh jumlah kelompok optimal yang terbentuk sebanyak tiga kelompok. Berdasarkan hasil pengelompokkan dengan VW K-Means diketahui bahwa variabel angka kematian kasar adalah variabel yang paling penting dalam pembentukan kelompok 1 dan 2. Sedangkan angka kematian kasar dan angka kematian balita adalah variabel yang paling penting dalam pembentukan kelompok 3. Selain itu, diketahui bahwa kelompok 1 terdiri dari 14 kabupaten/kota, kelompok 2 terdiri dari 19 kabupaten/kota, dan kelompok 3 terdiri dari 5 kabupaten/kota. Hasil penelitian berupa hasil pengelompokkan (klasterisasi) yang dapat digunakan untuk menggambarkan distribusi pengelompokkan kabupaten berdasarkan kondisi profil kesehatan daerah tersebut.

Kata kunci : *Mortalitas, Metode Silhouette, K-Means, VW K-Means.*

ABSTRACT

One indicator that describes health status, development and health services is the mortality rate. Mortality was formed from several indicators, namely crude birth rates, crude mortality, infant mortality, under-five mortality and maternal mortality. In 2014, the achievement of infant mortality was still above the agreed MDG target. The distribution of health indicators, mortality, in East Java Province needs to be known to achieve targets in the MDGs and to improve health status in East Java Province. This study discusses the grouping of regencies (and cities) in East Java Province based on the similarity of five mortality indicators using the Variable Weighted K-Means (VW K-Means) method. Based on the silhouette method, three optimal groups were formed. Based on the results of grouping with VW K-Means it is known that the crude mortality variable is the most important variable in the formation of groups 1 and 2. While the crude mortality rate and under-five mortality rate are the most important variables in group formation 3. In addition, it is known that groups 1 consists of 14 districts / cities, group 2 consists of 19 districts / cities, and group 3 consists of 5 districts / cities. The results of the study are in the form of grouping (clustering) that can be used to describe the distribution of district groups based on the condition of the health profile of the area.

Keywords : *Mortality, Silhouette Method, K-Means, VW K-Means*

1. PENDAHULUAN

MDGs atau Millennium Development Goals adalah sebuah deklarasi yang dilakukan oleh Pemerintah Indonesia Bersama negara PBB lainnya untuk mencapai kesejahteraan rakyat dan pembangunan masyarakat pada Tahun 2015. Bidang kesehatan merupakan salah satu indikator untuk mengetahui keberhasilan dari tingkat kesejahteraan dan pembangunan masyarakat di suatu negara. Menurut Dinkes Jatim (2015), angka kematian (mortalitas) merupakan salah satu indikator pembangunan kesehatan yang dapat menggambarkan situasi derajat kesehatan. Mortalitas di suatu wilayah dari waktu ke waktu dapat digunakan untuk memberikan gambaran perkembangan derajat kesehatan dan dasar penilaian keberhasilan program pembangunan dan pelayanan kesehatan.

Kemenkes RI (2010) menyatakan mortalitas dibentuk dari angka kelahiran kasar, angka kematian kasar, angka kematian bayi, angka kematian balita, dan angka kematian ibu. Kelima indikator tersebut menjadi target yang harus terus dioptimalkan untuk mencapai tujuan yang telah disepakati oleh Pemerintah Indonesia dalam perjanjian MDGs.

Menurut Data Dinas Kesehatan Provinsi Jatim, angka kematian ibu (AKI) Provinsi Jawa Timur mencapai 93,52 per 100.000 kelahiran hidup. Target AKI menurut MDGs yang harus dicapai pada tahun 2015 sebesar 102 per 100.000 kelahiran hidup. Artinya pada tahun 2014, Provinsi Jatim berhasil memenuhi target yang harus dicapai dalam MDGs. Sedangkan capaian Angka Kematian bayi (AKB) dan Angka Kematian Neonatal (AKN) di Provinsi Jawa Timur masih berada diatas target MDGs yang harus dicapai (Dinkes Jatim, 2015). Untuk mencapai target angka mortalitas yang harus dipenuhi dalam perjanjian MDGs dan untuk meningkatkan derajat kesehatan, diperlukan dukungan dari provinsi ke kabupaten/kota berupa fasilitas kesehatan dari berbagai aspek.

Cluster analysis atau analisis kelompok merupakan suatu metode statistika yang digunakan untuk mengklasifikasikan satu atau beberapa entitas atau individu yang mempunyai kesamaan atribut ke dalam suatu kelompok. Menurut Hung dkk. (2011), analisis kelompok adalah metode pengelompokan kumpulan data ke dalam kelompok yang menampilkan karakteristik yang serupa. Dengan kata lain, pengelompokan dilakukan berdasarkan kesamaan (similarity) atau jarak (dissimilarity).

K-Means merupakan salah satu metode non-hierarchical clustering yang dapat dikategorikan sebagai metode nonparametric. K-Means merupakan metode pengelompokan yang paling populer dan

menghasilkan satu entitas tepat diklasifikasikan ke dalam satu kelompok (tidak saling tumpang tindih atau overlap). Walaupun demikian, K-Means mempunyai beberapa kelemahan, yaitu tidak ada dasar untuk menentukan jumlah kelompok (cluster) yang optimal, menginisialisasi pusat cluster (centroid), dan menentukan variabel atau atribut utama dalam pembentukan cluster. Selain itu, K-Means tidak dapat memilih atau menentukan variabel yang paling besar kontribusinya dalam pembentukan cluster karena semua variabel diperlakukan sama dalam proses pengelompokan (clustering process). Oleh karena itu, untuk mengakomodasi kelemahan dalam K-Means, Huang dkk. (2005) mengembangkan algoritma K-Means dengan pemberian bobot pada setiap variabel, yaitu metode Variable Weighting K-Means (VW K-Means).

VW K-Means berguna untuk mengetahui variabel-variabel yang paling penting dalam pembentukan sebuah kelompok dengan menghitung bobot variabel dalam setiap iterasi. Variabel-variabel yang memiliki bobot paling rendah dapat dihilangkan dan pengelompokan dilakukan pada variabel yang tersisa untuk mencapai hasil pengelompokan yang baik (Baswade dkk., 2012). Berdasarkan penelitian oleh Huang dkk. (2005), akurasi pengelompokan yang dihasilkan oleh VW K-Means lebih besar daripada akurasi yang dihasilkan oleh K-Means. Dengan menggunakan algoritma VW K-Means, bobot yang dihasilkan mampu membedakan variabel yang paling penting dari seluruh variabel yang digunakan pada suatu kelompok. Huang dkk. (2005) menunjukkan bahwa variabel yang tidak signifikan dapat diidentifikasi sesuai dengan nilai bobot. Oleh karena itu, dapat dikatakan bahwa algoritma VW K-Means lebih baik daripada algoritma K-Means.

Beberapa penelitian telah dilakukan dengan menggunakan algoritma VW K-Means. Jing dkk., (2006) menggunakan VW K-Means untuk melakukan pengelompokan pada data teks (text clustering). Hung dkk. (2011) menerapkan algoritma VW K-Means pada color image segmentation. Penelitian terhadap VW K-Means dilakukan oleh Hung dkk. (2008) yang meneliti tentang variasi pada pendekatan VW K-Means.

Berdasarkan uraian diatas, penelitian ini dilakukan untuk mengetahui distribusi kondisi indikator kesehatan dengan mengelompokkan kabupaten-kabupaten (dan kota) di Provinsi Jawa Timur berdasarkan kemiripan lima indikator mortalitas. Pengelompokan dilakukan dengan metode VW K-Means. Algoritma VW K-Means digunakan karena mampu memberikan akurasi atau

ketepatan pengelompokan yang tinggi. Selain itu, dengan menggunakan VW K-Means, variabel indikator mortalitas yang paling penting dalam kelompok dapat diketahui. Variabel yang paling penting dalam kelompok yang terbentuk yang dapat digunakan sebagai informasi dasar untuk mengoptimalkan derajat kesehatan mortalitas di Provinsi Jawa Timur.

2. METODE PENELITIAN

2.1 Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini berupa data sekunder yang diperoleh dari Profil Kesehatan Provinsi Jawa Timur 2014 oleh Dinkes Provinsi Jawa Timur. Variabel yang digunakan dalam penelitian ini meliputi data jumlah penduduk, jumlah kematian, jumlah lahir hidup, jumlah kematian bayi, jumlah penduduk balita, jumlah kematian balita, dan jumlah kematian ibu. Ketujuh data tersebut digunakan untuk menghitung nilai variabel penelitian, yaitu angka kematian kasar (x1), angka kelahiran kasar (x2), angka kematian bayi (x3), angka kematian balita (x4), dan angka kematian ibu (x5).

2.2 Metode Analisis Data

Metode yang digunakan adalah Variable Weighting K-Means dengan pemilihan jumlah kelompok dengan menggunakan metode Silhouette. Adapun langkah analisis adalah sebagai berikut.

1. Mendeskripsikan setiap variabel penelitian (x1-x5) dengan ukuran pemusatan dan penyebaran data.
2. Menentukan jumlah kelompok terbaik dengan metode *Silhouette*, dicobakan 3 sampai 7 kelompok.

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)} \tag{1}$$

dimana

- ai : jarak rata-rata observasi i ke observasi lain yang ada dalam kelompok yang sama
- bi : jarak terkecil antar d(i, C), atau d(i, C) : jarak antara observasi i ke kelompok C yang bukan merupakan kelompok observasi i berada.

3. Mengelompokkan kabupaten/kota yang ada di Jawa Timur berdasarkan kesamaan 5 variabel (x1-x5) dengan metode *VW K-Means* dengan langkah-langkah sebagai berikut:

1. Menentukan sejumlah k kelompok dengan menggunakan jumlah kelompok optimal yang terbentuk pada langkah (2).

2. Menentukan inisialisasi centroid $Z^{(0)}$ kelompok secara acak pada rentang (minimum dan maksimal) nilai setiap variabel.
3. Menentukan nilai parameter bobot $\beta = 10$ (parameter yang dihasilkan oleh akurasi tertinggi pada Huang dkk. (2005)).

$$P(U, Z, W) = \sum_{l=1}^k \sum_{i=1}^n \sum_{j=1}^m u_{i,l} w_j^\beta d(x_{i,j}, z_{l,j}) \tag{2}$$

Dengan

$$\begin{cases} \sum_{l=1}^k u_{i,l} = 1, & 1 \leq i \leq n \\ u_{i,l} \in \{0,1\}, & 1 \leq i \leq n, 1 \leq l \leq k \\ \sum_{j=1}^m w_j = 1, & 0 \leq w_j \leq 1 \end{cases}$$

4. Menginisialisasi bobot kelima variabel, $w^0 = [w_{x_1}^0, w_{x_2}^0, \dots, w_{x_5}^0]$, dimana nilai awal untuk elemen adalah $1/m = 1/5$, sehingga $\sum_{j=1}^5 w_j = 1$.

5. Untuk iterasi pertama, hitung $U^{(h+1)}$, hitung jarak variabel yang dimiliki oleh kabupaten/kota ke masing-masing centroid *cluster* yang telah ditetapkan pada langkah (2) dengan bobot yang telah ditetapkan pada langkah (4). Bandingkan jarak observasi ke setiap kelompok. Masukkan kabupaten/kota menjadi anggota kelompok yang memiliki jarak terkecil.

$$\begin{cases} u_{i,l} = 1, & \text{jika } \sum_{j=1}^m w_j^\beta d(x_{i,j}, z_{l,j}) \leq \sum_{j=1}^m w_j^\beta d(x_{i,j}, z_{t,j}) \text{ , untuk } 1 \leq t \leq k \\ u_{i,t} = 0, & \text{, untuk } t \neq l \end{cases} \tag{3}$$

6. Untuk iterasi pertama, hitung $Z^{(h+1)}$, dimana $Z^{(h+1)}$ diperoleh sebagai berikut.

$$Z_{l,j} = \frac{\sum_{i=1}^n u_{i,l} x_{i,j}}{\sum_{i=1}^n u_{i,l}} \text{ , untuk } 1 \leq l \leq k \tag{4}$$

dan $1 \leq j \leq m$

Hitung $P(U^{(h+1)}, Z^{(h+1)}, W^h)$. Jika $P(U^{(h+1)}, Z^{(h+1)}, W^h) = P(U^{(h+1)}, Z^h, W^h)$

maka hentikan perhitungan, jika tidak lanjutkan ke step ketujuh.

7. Untuk iterasi pertama, hitung $W^{(h+1)}$.

$$\hat{w}_j = \begin{cases} 0, & \text{jika } D_j = 0 \\ \frac{1}{\sum_{t=1}^h \left[\frac{D_j}{D_t} \right]^{\beta-1}}, & \text{jika } D_j \neq 0 \end{cases} \quad (5)$$

Jika

$$P(U^{(h+1)}, Z^{(h+1)}, W^{(h+1)}) = P(U^{(h+1)}, Z^{(h+1)}, W^h)$$

maka hentikan perhitungan, jika tidak, tetapkan $h = h + 1$ dan lakukan iterasi kedua dengan tahapan yang sama seperti step kelima hingga ketujuh.

8. Memetakan hasil pengelompokkan dengan *VW K-Means*.

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi Indikator Mortalitas

Rata-rata angka kematian kasar (x_1) di Provinsi Jawa Timur yang terjadi pada tahun 2014 adalah sebesar 0.1036, artinya terdapat 1 kematian (0.1036 dibulatkan keatas) untuk tiap 1000 penduduk. Angka kematian kasar terendah terdapat di Kabupaten Bangkalan dengan nilai sebesar 0.0200 sedangkan angka tertinggi terdapat di Kota Madiun 0.2980. Varians data dari angka kematian kasar sebesar 0.0033 (Tabel 1).

Tabel 1. Deskripsi Lima Indikator Mortalitas

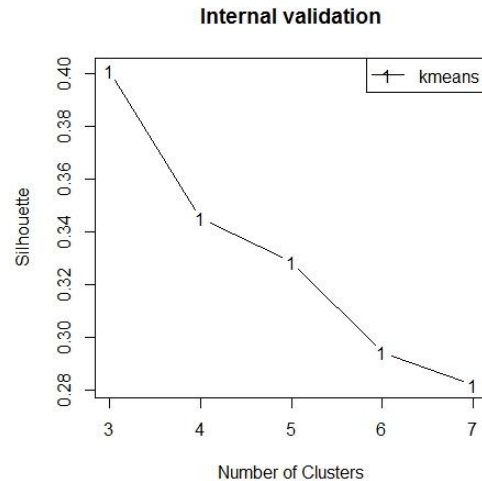
Nilai	x_1	x_2	x_3
Min	0.020	12.321	3.348
Max	0.298	39.398	24.588
Rata-rata	0.104	16.202	9.303
Varians	0.003	16.961	18.957
Nilai	x_4	x_5	
Min	0.840	0.291	
Max	5.286	2.093	
Rata-rata	2.065	0.983	
Varians	0.842	0.150	

Angka kelahiran kasar di Kabupaten Pacitan sebesar 12.3210. Artinya, terdapat 13 (12.3210 dibulatkan keatas) kelahiran per 1000 penduduk di Kabupaten Pacitan. Angka tersebut merupakan angka terendah dari seluruh kabupaten/kota lainnya di Provisini Jawa Timur. Sedangkan angka kelahiran tertinggi berada di Kota Madiun, yaitu sebesar 39.3980. Rata-rata angka kelahiran

kasar yang terjadi di Jawa Timur sebesar 16.2019 (Tabel 1).

3.2 Penentuan Jumlah Kelompok dengan Metode *Silhouette*

Gambar 1 menunjukkan dengan metode *Silhouette*, lebar *Silhouette* tertinggi berada pada jumlah kelompok 3, dengan nilai sebesar 0.4008.



Gambar 1. Jumlah Kelompok berdasarkan Lebar *Silhouette*

Oleh sebab itu, jumlah kelompok yang akan dibentuk untuk mengelompokkan kabupaten/kota di Jawa Timur berdasarkan kemiripan indikator mortalitas adalah 3 kelompok.

Tabel 2. Lebar *Silhouette*

Jumlah Kelompok				
3	4	5	6	7
*0.4008	0.3450	0.3286	0.2942	0.2820

Keterangan: *) Jumlah Kelompok Optimum

3.3. Pengelompokkan Kabupaten/Kota yang Ada di Jawa Timur berdasarkan Kesamaan 5 Variabel (x_1 - x_5) dengan Metode *VW K-Means*

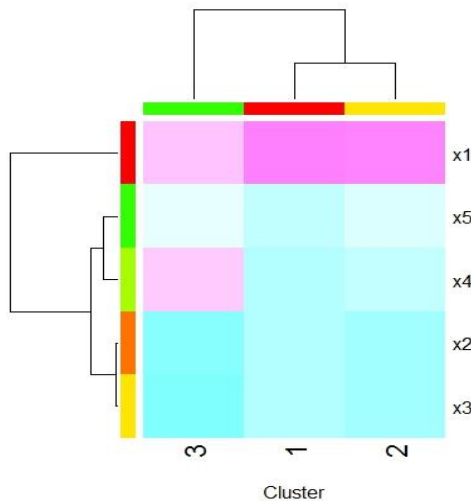
Berdasarkan penentuan jumlah kelompok dengan metode *Silhouette*, diketahui bahwa jumlah kelompok optimum yang terbentuk adalah 3 kelompok. Jumlah anggota dari tiap kelompok dengan menggunakan *VW K-Means* ditampilkan

pada Tabel 3. Kelompok 1 terdiri dari 14 kabupaten/kota yang memiliki kesamaan lima indikator mortalitas.

Tabel 3. Jumlah Keanggotaan Tiap Kelompok dengan *VW K-Means*

Kelompok 1	Kelompok 2	Kelompok 3
14	19	5

Gambar 2 menunjukkan tingkat kepentingan variabel dalam kelompok. Variabel yang mempunyai warna gelap adalah variabel yang memiliki bobot tinggi dalam kelompok. Variabel dengan bobot tertinggi atau variabel yang paling penting dalam kelompok 1 adalah variabel x_1 , dengan nilai sebesar 0.9364. Variabel x_1 juga merupakan variabel yang paling penting dalam membentuk kelompok 2, dimana nilai bobot untuk variabel x_1 sebesar 0.7041.



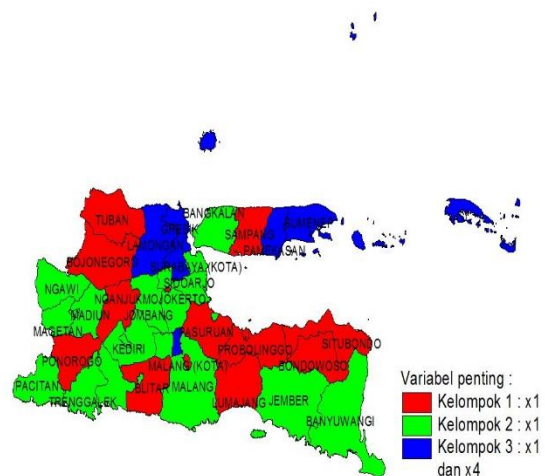
Gambar 2. Visualisasi Tingkat Kepentingan Variabel di Tiap Kelompok

Sedangkan variabel terpenting dalam kelompok 3 adalah variabel x_1 dan x_4 dengan nilai masing-masing 0.3395 dan 0.3277 (Tabel 4). Berdasarkan perhitungan dengan *VW K-Means* diketahui bahwa Kabupaten Ponorogo, Kabupaten Lumajang, dan Kabupaten Bondowoso merupakan anggota dari kelompok 1, dimana variabel penting dalam kelompok adalah angka kematian kasar (x_1). Kabupaten Pacitan, Kabupaten Trenggalek, dan Kabupaten Tulungagung termasuk dalam kelompok 2.

Tabel 4. Nilai Bobot Variabel untuk Tiap Kelompok

Kelompok	x_1	x_2	x_3
1	0.9364	0.0000	0.0000
2	0.7041	0.0000	0.0000
3	0.3395	0.0675	0.0532
Kelompok	x_4	x_5	
1	0.0002	0.0633	
2	0.1065	0.1894	
3	0.3277	0.2122	

Sedangkan beberapa anggota kelompok dalam kelompok ketiga adalah Kabupaten Lamongan, Kabupaten Gresik, dan Kabupaten Pamekasan (Gambar 3).



Gambar 3. Hasil Pengelompokan dengan *VW K-Means*

Tabel 5 merupakan *centroid* atau nilai rata-rata variabel dari masing-masing kelompok yang dapat digunakan sebagai perbandingan indikator antar kelompok. Kelompok yang terbentuk sebanyak 3 kelompok, perbedaan nilai dari ketiga

kelompok tersebut dapat membedakan kelompok dengan tingkat tinggi, sedang, dan rendah. Kelompok pertama merupakan kabupaten/kota yang memiliki kemiripan variabel angka kematian kasar (x_1), angka kematian bayi (x_3), angka kematian balita (x_4), dan angka kematian ibu (x_5) dengan tingkat tinggi, sedangkan angka kelahiran kasar (x_2) yang rendah.

Tabel 5. Nilai Rata-rata (*Centroid*) Variabel tiap Kelompok

Kelompok	x_1	x_2	x_3
1	0.1270 (Tinggi)	15.5283 (Rendah)	13.7561 (Tinggi)
2	0.0961 (Sedang)	16.8285 (Tinggi)	7.3242 (Sedang)
3	0.0664 (Rendah)	15.7068 (Sedang)	4.3562 (Rendah)
Kelompok	x_4	x_5	
1	2.9901 (Tinggi)	1.1775 (Tinggi)	
2	1.6662 (Sedang)	0.9074 (Sedang)	
3	0.9862 (Rendah)	0.7244 (Rendah)	

4. KESIMPULAN DAN SARAN

Kesimpulan yang diperoleh dari penelitian ini adalah dengan menggunakan VW K-Means diketahui bahwa dari tiga kelompok yang dibentuk. Kelompok pertama dapat digolongkan sebagai kelompok yang mempunyai nilai tinggi, kelompok kedua adalah kelompok dengan sedang, dan kelompok ketiga adalah kelompok yang mempunyai nilai rendah. Berdasarkan hasil pembobotan, kelompok pertama merupakan kabupaten/kota yang memiliki kemiripan variabel angka kematian kasar (x_1), angka kematian bayi (x_3), angka kematian balita (x_4), dan angka kematian ibu (x_5) dengan tingkat tinggi, sedangkan angka kelahiran kasar (x_2) yang rendah. Variabel angka kematian kasar (x_1) adalah variabel yang paling penting dalam membentuk kelompok pertama. Berdasarkan hasil pengelompokan dengan VW K-Means, diketahui bahwa kelompok yang ada dalam satu kelompok sebagian besar berdekatan secara lokasi.

DAFTAR PUSTAKA

Baswade, A.M., Joshi, K.D., dan Nalwade, P.S. 2012. *A Comparative Study of K-Means and Weighted K-Means for Clustering*, *International Journal of Engineering*

Research & Technology (IJERT), Vol. 1 Issue 10, Desember 2012.

Dinas Kesehatan Provinsi Jatim [Dinkes Provinsi Jatim]. 2015. *Profil Kesehatan Provinsi Jawa Timur 2014*. Surabaya. Dinas Kesehatan Provinsi Jawa Timur.

Huang, J.Z., Ng, M.K., Rong, H., dan Li, Z. 2005. *Automated Variable Weighting in k-Means Type Clustering*, *Journal of IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27 No. 5, May 2005, pp 657-668.

Hung, W.L., Chang, Y.C., dan Lee, E.S. 2011. *Weight Selection In W-K-Means Algorithm With An Application In Color Image Segmentation*, *Journal of Computers and Mathematics with Applications*, Vol. 62, Issue 2, Juli 2011, pp 668-676.

Hung, W.L., Yang, M.S., Chen, D.H. 2008. *Variation Approaches to Feature-Weight Selection and Application to Fuzzy Clustering*, *IEEE International Conference on Fuzzy Systems*, pp 276-280.

Jing, L., Ng, M.K., Yang, X., dan Huang, J.Z. 2006. *A Text Clustering System based on K-Means Type Subspace Clustering and Ontology*, *International Journal of Intelligent Technology*, Vol. 1 No. 2.

Kaufman, L. dan Rousseeuw, P.J. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York. Wiley.

Kementerian Kesehatan Republik Indonesia [Kemenkes RI]. 2010. *Indeks Pembangunan Kesehatan Masyarakat IPKM*. Jakarta. Kementerian Kesehatan Republik Indonesia.

Sharma, S. 1996. *Applied Multivariate Techniques*. Canada. John Wiley & Sons, Inc.