

**PENGELOMPOKAN TINGKAT KEAMANAN WILAYAH JAWA
TENGAH BERDASARKAN INDEKS KEJAHATAN DAN JUMLAH
POS KEAMANAN DENGAN METODE KLASTERING K-MEANS**

Retno Tri Vulandari
Program Studi Sistem Informasi, STMIKSinarNusantaraSurakarta
retno.tv@gmail.com

Abstract

The level of security in Central Java declined but for each region is not the case. Based on the crime index and the number of security posts in each police region of Central Java. These data are used for grouping the two regions in Central Java, namely good security level and bad security level. K-Means clustering method is a method used to classify data based on multiple attributes. K-Means clustering method is a method used to classify data based on multiple attributes. In this study, will be applied to the K-Means method with random initial centers. It is derived from two police regions that describe the good and bad security level. The formulation of the problem is to know the areas with good and bad security level in Central Java. Based on the results, there are 8 areas with good security level and 27 areas with bad security level.

Keyword : Crime Index, K-Means Clustering

I. PENDAHULUAN

Kebutuhan dasar setiap manusia terdiri dari kebutuhan biologis seperti makan, minum serta tidur, dan kebutuhan sosial, seperti status sosial, peranan sosial, aktualisasi diri dan rasa aman. Saat ini dapat dikatakan bahwa rasa aman merupakan salah satu kebutuhan dasar manusia dalam menjalankan aktivitas sehari-harinya. Dalam teori hirarki kebutuhan manusia, rasa aman berada pada tingkatan yang kedua dibawah kebutuhan dasar manusia seperti sandang, pangan, dan papan. Hal ini menunjukkan bahwa rasa aman merupakan kebutuhan manusia yang penting (Maslow, 1984).

Rasa aman (*security*) merupakan salah satu hak asasi yang harus diperoleh atau dinikmati setiap orang. Rasa aman merupakan variabel yang sangat luas karena mencakup berbagai aspek dan dimensi, mulai dari dimensi politik, hukum, pertahanan, keamanan, sosial dan ekonomi. Sejalan dengan itu, statistik dan indikator yang biasa digunakan untuk mengukur rasa aman masyarakat merupakan indikator negatif, misalnya jumlah

angka kejahatan (*crime total*), jumlah orang yang berisiko terkena tindak kejahatan (*crime rate*) setiap 100.000 penduduk. Semakin tinggi angka kriminalitas menunjukkan semakin banyak tindak kejahatan pada masyarakat yang merupakan indikasi bahwa masyarakat merasa semakin tidak aman (Badan Pusat Statistika, 2014).

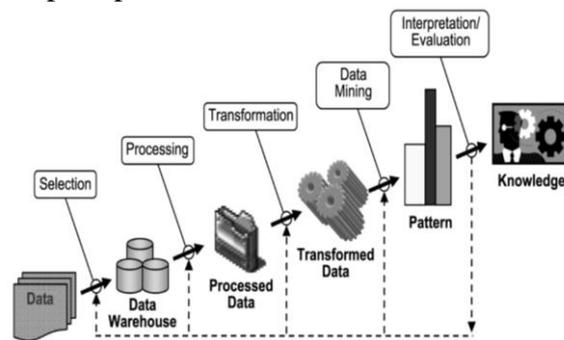
Oleh karena itu, akan diteliti mengenai pengelompokan tingkat keamanan wilayah daerah Jawa Tengah berdasarkan data indeks kejahatan dan jumlah pos keamanan. Batasan masalah dalam penulisan ini adalah menggunakan data indeks kejahatan dan jumlah pos keamanan di 35 wilayah kepolisian daerah Jawa Tengah tahun 2013.

II. TINJAUAN PUSTAKA

2.1 Data Mining

Data mining adalah suatu metode pengolahan data untuk menemukan pola yang tersembunyi dari data tersebut. Hasil dari pengolahan data dengan metode data mining ini dapat digunakan untuk mengambil keputusan di masa depan. Data mining ini juga dikenal dengan istilah *pattern recognition*. *Data mining* merupakan metode pengolahan data berskala besar oleh karena itu data mining ini memiliki peranan penting dalam bidang industri, keuangan, cuaca, ilmu dan teknologi. Secara umum kajian data mining membahas metode-metode seperti, clustering, klasifikasi, regresi, seleksi variabel, dan *market basket* analisis (Santosa, 2007).

Data mining merupakan inti dari proses *Knowledge Discovery in Database* (KDD), meliputi dugaan algoritma yang mengeksplor data, membangun model dan menemukan pola yang belum diketahui. KDD merupakan penyelesaian masalah dengan menganalisa data yang ada pada database dengan data tersimpan secara elektronik dan pencariannya dilakukan otomatis seperti pada komputer. Tahapan pada proses KDD pada database seperti pada Gambar 1.



Gambar 1. Tahapan Proses KDD

Sumber: (Agusta, 2007)

2.2 Metode Klaster

Pada dasarnya metode klaster merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. Klaster merupakan salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*). Ada dua jenis metode klaster yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering* (Santosa, 2007).

Hierarchical clustering adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian seterusnya sehingga klasterakan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah klaster. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut. Berbeda dengan metode *hierarchical clustering*, metode *non-hierarchical clustering* justru dimulai dengan menentukan terlebih dahulu jumlah klaster yang diinginkan. Setelah jumlah klaster diketahui, baru proses cluster dilakukan tanpa mengikuti proses hierarki (Santoso, 2010).

2.3 Metode Klaster K-Means

Metode klaster K-Means merupakan algoritma *clustering* yang berulang-ulang. Algoritma K-Means menetapkan nilai-nilai klaster (K) secara random, untuk sementara nilai tersebut menjadi pusat dari klaster atau biasa disebut dengan centroid, mean atau “means”. Kemudian menghitung jarak setiap data yang ada terhadap masing-masing centroid menggunakan rumus Euclidian hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid (Rismawan, 2008).

Klasifikasikan setiap data berdasarkan kedekatannya dengan centroid. Lakukan langkah tersebut hingga nilai centroid tidak berubah (stabil). Langkah – langkah metode K-Means (Santosa, 2007) adalah sebagai berikut

- a. menentukan pusat klaster awal secara acak
- b. alokasikan semua data obyek kedalam klaster dengan jarak terdekat. Demikian juga kedekatan suatu data ke klaster tertentu ditentukan jarak antara data dengan pusat klaster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat klaster. Jarak antara satu data dengan satu pusat klaster tertentu akan menentukan suatu data masuk

dalam kluster mana. Untuk menghitung jarak semua data ke setiap titik pusat kluster dapat menggunakan teori jarak Euclidean yang dirumuskan (Badan Pusat Statistika, 2014) sebagai berikut:

$$d(i, j) = \sqrt{(x_{1i} - x_{1j})^2 + \dots + (x_{ki} - x_{kj})^2} \quad (1)$$

dengan

- $d(i, j)$: jarak data kekepusat kluster j
- x_{ki} : data ke i pada atribut data ke k
- x_{kj} : titik pusat ke j pada atribut ke k

- c. Hitung kembali pusat kluster dengan keanggotaan kluster yang sekarang. Pusat kluster adalah rata-rata dari semua data dalam kluster tertentu. Jika dikehendaki bisa juga menggunakan median dari kluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
- d. Tugaskan lagi setiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai atau kembali ke langkah sebelumnya sampai pusat cluster tidak berubah lagi.

2.4 Indeks Kejahatan

Indeks kejahatan (*Crime Index*) yaitu jenis-jenis kejahatan yang digunakan sebagai alat pengukur dalam statistik kriminal seperti:

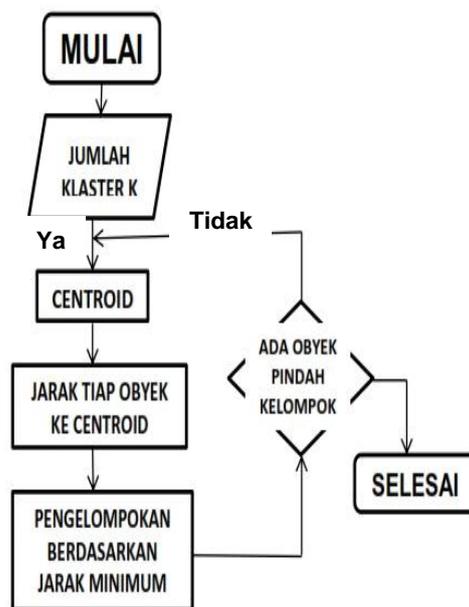
- a. Kejahatan tersebut dianggap sebagai kejahatan yang serius oleh masyarakat
- b. Frekuensi terjadinya kejahatan tersebut cukup besar atau cukup sering
Jadi indeks kejahatan itu tidak sama untuk semua tempat atau wilayah. Contoh indeks kejahatan Polda Jawa Tengah 1978: pembakaran dan kebakaran, kejahatan terhadap mata uang, pembunuhan, penganiayaan berat, pencurian dengan pemberatan, pencurian dengan kekerasan, pencurian kendaraan bermotor, dan penyalahgunaan narkotik (Badan Pusat Statistika, 2014). Indeks kejahatan (I_t) dirumuskan sebagai berikut

$$I_t = \frac{\text{Jumlah kejahatan pada tahun } t}{\text{Jumlah kejahatan pada tahun dasar}} \times 100 \quad (2)$$

III. METODE PENELITIAN

3.1 Gambaran Umum Model

Pada penelitian ini akan dibangun sistem yang dapat digunakan untuk mengelompokkan tingkat keamanan wilayah Jawa Tengah berdasarkan indeks kejahatan dan jumlah pos keamanan. Kemudian langkah selanjutnya adalah melakukan proses pengelompokan data menggunakan metode klaster K-Means. Pengelompokan menjadi dua kelompok yang merupakan wilayah dengan tingkat keamanan baik dan buruk, dengan pusat awal mengambil dua wilayah kepolisian di Jawa Tengah yang menggambarkan kedua keadaan tersebut. Adapun diagram alir proses klastering K-means dapat dilihat pada Gambar 2.



Gambar 2. Diagram Alir K-Means

3.2 Data

Data indeks kejahatan dan jumlah pos keamanan merupakan data sekunder yang diperoleh dari www.bps.go.id tahun 2013. Data-data tersebut dapat dilihat pada Tabel 1.

Tabel 1..Data Indeks Kejahatan dan Pos Keamanan

Wilayah	Indek Jahat	Pos Aman
Res. Cilacap	177	7,054
Res. Banyumas	239	4,784
Res. Purbalingga	144	2,494
Res. Banjarnegara	68	1,927

Tabel 1. Lanjutan

Wilayah	Indek Jahat	Pos Aman
Res. Kebumen	123	3,052
Res. Purworejo	102	2,258
Res. Wonosobo	69	1,256
Res. Magelang	121	2,835
Res. Boyolali	108	3,043
Res. Klaten	213	1,418
Res. Sukoharjo	170	5,019
Res. Wonogiri	44	3,566
Res. Karanganyar	126	3,719
Res. Sragen	112	4,508
Res. Grobogan	87	2,790
Res. Blora	61	670
Res. Rembang	79	1,396
Res. Pati	92	4,338
Res. Kudus	76	866
Res. Jepara	127	1,370
Res. Demak	116	2,826
Res. Semarang	147	3,504
Res. Temanggung	94	1,381
Res. Kendal	131	2,287
Res. Batang	99	1,718
Res. Pekalongan	70	1,016
Res. Pemasang	77	2,116
Res. Tegal	107	1,346
Res. Brebes	93	1,247
ResTa. Magelang	43	207
ResTa. Surakarta	407	1,498
Res. Salatiga	143	719
ResTabes. Semarang	560	1,824
ResTa. Pekalongan	116	514
ResTa. Tegal	84	160

Sumber: BPS Jawa Tengah 2013

IV. HASIL DAN PEMBAHASAN

Adapun langkah dari pengelompokan data adalah sebagai berikut

- a. Tentukan pusat kluster secara acak, dalam penelitian ini diambil pusat kluster awal, wilayah Klaten sebagai wilayah tingkat keamanan baik $c_1 = (213, 1.418)$ dan wilayah Wonogiri sebagai wilayah tingkat keamanan buruk $c_2 = (44, 3.566)$
- b. Hitung jarak setiap data yang ada terhadap setiap pusat kluster. Misalkan untuk menghitung jarak data pertama dengan pusat kluster pertama adalah :

$$d_{11} = \sqrt{(177 - 213)^2 + (7054 - 1418)^2}$$

$$= 5.636,115$$

Jarak data mahasiswa pertama dengan pusat kluster kedua:

$$d_{12} = \sqrt{(177 - 44)^2 + (7054 - 3566)^2}$$

$$= 3.490,535$$

Jarak data mahasiswa kedua dengan pusat kluster pertama:

$$d_{21} = \sqrt{(239 - 213)^2 + (4784 - 1418)^2}$$

$$= 5.636,115$$

Jarak data mahasiswa kedua dengan pusat kluster kedua:

$$d_{22} = \sqrt{(239 - 44)^2 + (4784 - 3566)^2}$$

$$= 3.490,535$$

dan seterusnya seperti pada Tabel 2.

Tabel 2. Hasil Perhitungan Jarak Kluster

Wilayah	C1	C2
Res. Cilacap	5636.115	3490.535
Res. Banyumas	3366.1	1233.511
Res. Purbalingga	1078.21	1076.654
Res. Banjarnegara	529.2504	1639.176
Res. Kebumen	1636.477	520.0356
Res. Purworejo	847.3022	1309.285
Res. Wonosobo	216.7487	2310.135
Res. Magelang	1419.983	735.0442
Res. Boyolali	1628.389	526.9013

Tabel 2. Lanjutan

Wilayah	C1	C2
Res. Klaten	0	2154.638
Res. Sukoharjo	3601.257	1458.453
Res. Wonogiri	2154.638	0
Res. Karanganyar	2302.644	173.5886
Res. Sragen	3091.65	944.4512
Res. Grobogan	1377.774	777.1905
Res. Blora	763.2876	2896.05
Res. Rembang	135.794	2170.282
Res. Pati	2922.506	773.4908
Res. Kudus	568.7469	2700.19
Res. Jepara	98.48858	2197.568
Res. Demak	1411.337	743.4945
Res. Semarang	2087.044	120.2206
Res. Temanggung	124.6194	2185.572
Res. Kendal	872.8602	1281.956
Res. Batang	320.9299	1848.818
Res. Pekalongan	426.6767	2550.133
Res. Pemasang	711.1259	1450.375
Res. Tegal	128.1405	2220.894
Res. Brebes	208.9043	2319.518
ResTa. Magelang	1222.874	3359
ResTa. Surakarta	209.8476	2099.617
Res. Salatiga	702.4963	2848.721
ResTabes. Semarang	534.0833	1816.816
ResTa. Pekalongan	909.1892	3052.849
ResTa. Tegal	1264.597	3406.235

- c. Suatu data akan menjadi anggota dari suatu kluster yang memiliki jarak terkecil dari pusat klusternya. Misal untuk data pertama, jarak terkecil diperoleh pada kluster kedua, sehingga data pertama akan menjadi anggota dari kluster kedua. Posisi kluster selengkapnya dapat dilihat pada Tabel 3.

Tabel 3. Posisi Klaster pada Iterasi Pertama

Wilayah	C1	C2
Res. Cilacap		*
Res. Banyumas		*
Res. Purbalingga		*
Res. Banjarnegara	*	
Res. Kebumen		*
Res. Purworejo	*	
Res. Wonosobo	*	
Res. Magelang		*
Res. Boyolali		*
Res. Klaten	*	
Res. Sukoharjo		*
Res. Wonogiri		*
Res. Karanganyar		*
Res. Sragen		*
Res. Grobogan		*
Res. Blora	*	
Res. Rembang	*	
Res. Pati		*
Res. Kudus	*	
Res. Jepara	*	
Res. Demak		*
Res. Semarang		*
Res. Temanggung	*	
Res. Kendal	*	
Res. Batang	*	
Res. Pekalongan	*	
Res. Pemalang	*	
Res. Tegal	*	
Res. Brebes	*	
ResTa. Magelang	*	
ResTa. Surakarta	*	
Res. Salatiga	*	

Tabel 3. Lanjutan

Wilayah	C1	C2
ResTabes. Semarang	*	
ResTa. Pekalongan	*	
ResTa. Tegal	*	

- d. Hitung pusat kluster baru. Untuk kluster pertama, ada 22 sehingga:

$$c_{11} = \frac{(529,25 + \dots + 1264,59)}{22} = 514,09$$

$$c_{12} = \frac{(1639,17 + \dots + 3406,23)}{22} = 2.277,04$$

Untuk kluster kedua, ada 13 data

$$c_{21} = \frac{(5636,12 + \dots + 2087,04)}{22} = 2408,15$$

$$c_{22} = \frac{(3490,54 + \dots + 120,22)}{22} = 898,11$$

- e. Ulangi langkah 2 hingga posisi data sudah tidak mengalami perubahan. Dalam penelitian ini pada iterasi ke-9 dan iterasi ke-10 tidak terjadi perubahan. Kelompok atau kluster pertama terdapat 8 anggota dan kluster kedua terdapat 27 anggota.

Tabel 4. Posisi Kluster pada Iterasi ke-10

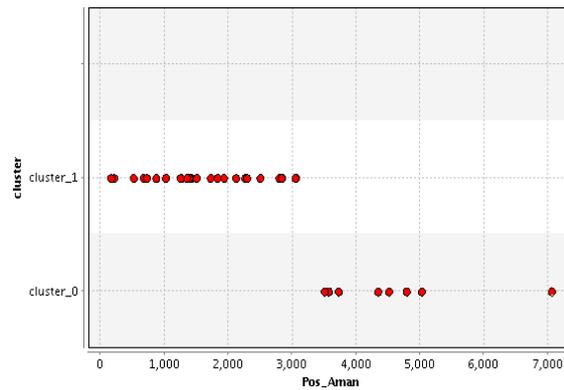
Wilayah	C1	C2
Res. Cilacap		*
Res. Banyumas		*
Res. Purbalingga	*	
Res. Banjarnegara	*	
Res. Kebumen	*	
Res. Purworejo	*	
Res. Wonosobo	*	
Res. Magelang	*	
Res. Boyolali	*	
Res. Klaten	*	

Tabel 4. Lanjutan

Wilayah	C1	C2
Res. Sukoharjo		*
Res. Wonogiri		*
Res. Karanganyar		*
Res. Sragen		*
Res. Grobogan	*	
Res. Blora	*	
Res. Rembang	*	
Res. Pati		*
Res. Kudus	*	
Res. Jepara	*	
Res. Demak	*	
Res. Semarang		*
Res. Temanggung	*	
Res. Kendal	*	
Res. Batang	*	
Res. Pekalongan	*	
Res. Pemalang	*	
Res. Tegal	*	
Res. Brebes	*	
ResTa. Magelang	*	
ResTa. Surakarta	*	
Res. Salatiga	*	
ResTabes. Semarang	*	
ResTa. Pekalongan	*	
ResTa. Tegal	*	

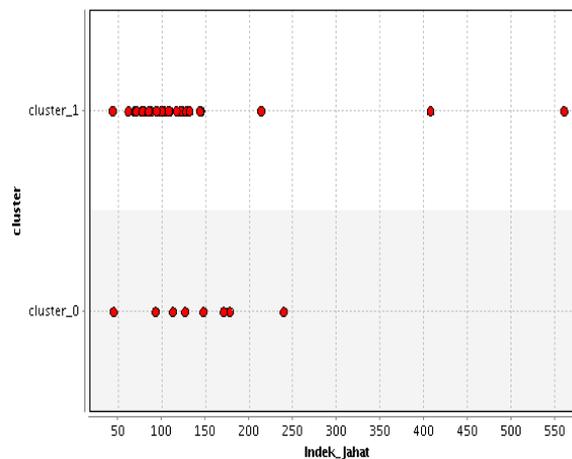
Berdasarkan Tabel 4, wilayah yang mempunyai tingkat keamanan buruk terdapat 27 wilayah, yaitu: Purbalingga, Banjarnegara, Kebumen, Purworejo, Wonosobo, Magelang, Boyolali, Klaten, Grobogan, Blora, Rembang, Kudus, Jepara, Demak, Temanggung, Kendal, Batang, Pekalongan, Pemalang, Tegal, Brebes, Polresta Magelang, Surakarta, Salatiga, Semarang, Pekalongan, dan Tegal. Wilayah yang mempunyai tingkat keamanan baik terdapat 8 wilayah, yaitu: Cilacap, Banyumas,

Sukoharjo, Wonogiri, Karanganyar, Sragen, Pati, dan Semarang. Hal tersebut dapat dilihat pada Gambar 3, Gambar 4, dan Gambar 5.

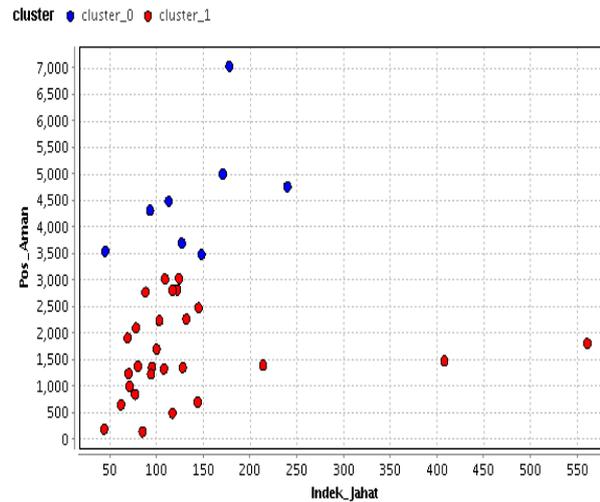


Gambar 3. Pengelompokan berdasarkan Jumlah Pos Keamanan

Berdasarkan Gambar 3, terlihat jika jumlah pos keamanan lebih dari 3.000 maka akan termasuk dalam klaster pertama atau tingkat keamanan baik, dan sebaliknya jika jumlah pos keamanan kurang dari 3.000 akan termasuk dalam klaster kedua atau tingkat keamanan buruk. Berdasarkan Gambar 4, tidak dapat diketahui dengan pasti pengaruh nilai indeks kejahatan terhadap tingkat keamanan karena daerah dengan indeks kejahatan rendah ada juga yang masuk dalam kelompok wilayah tingkat keamanan buruk. Oleh karena itu, proses pengelompokan diperlukan dua faktor, yaitu indeks kejahatan dan jumlah pos keamanan seperti terlihat pada Gambar 5.



Gambar 4. Pengelompokan berdasarkan Indeks Kejahatan



Gambar 5. Pengelompokan berdasarkan Jumlah Pos Keamanan dan Indeks Kejahatan

V. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil penelitian, dapat disimpulkan bahwa algoritma klustering K-Means dapat digunakan untuk mengelompokkan tingkat keamanan wilayah daerah Jawa Tengah berdasarkan indeks kejahatan dan jumlah pos keamanan 2013. Terdapat 8 wilayah dengan tingkat keamanan baik, yaitu: Cilacap, Banyumas, Sukoharjo, Wonogiri, Karanganyar, Sragen, Pati, dan Semarang. Terdapat 27 wilayah dengan tingkat keamanan buruk, yaitu Purbalingga, Banjarnegara, Kebumen, Purworejo, Wonosobo, Magelang, Boyolali, Klaten, Grobogan, Blora, Rembang, Kudus, Jepara, Demak, Temanggung, Kendal, Batang, Pekalongan, Pemalang, Tegal, Brebes, Polresta Magelang, Surakarta, Salatiga, Semarang, Pekalongan, dan Tegal.

5.2 Saran

Penelitian ini merupakan pengelompokan tingkat keamanan wilayah Jawa Tengah yang dipengaruhi oleh dua faktor yaitu indeks kejahatan dan jumlah pos keamanan, untuk penelitian selanjutnya dapat dikembangkan pengelompokan tingkat keamanan dengan dipengaruhi faktor lain seperti jumlah aparat keamanan, jumlah CCTV kota, pendapatan daerah, tingkat kemiskinan, dan sebagainya. Untuk penelitian selanjutnya dapat dikembangkan dengan metode kluster lain seperti KNN, *Mixture Modelling*, *Self-Organising Map* (SOM), Algoritma *Rock*, dan sebagainya.

DAFTAR PUSTAKA

- Agusta, Yudi. 2007. K-Means-Penerapan, Permasalahan, dan Metode Terkait. *Jurnal Sistem dan Informatika*. Vol 3: 47 – 60.
- Badan Pusat Statistika. 2014. *Statistik Kriminalitas 2014*. Jakarta: BPS Press.
- Maslow, Abraham, H. 1984. *Motivasi dan Kepribadian : Teori Motivasi dengan Rancangan Hirarki Kebutuhan Manusia*. Jakarta : Pustaka Binaman Pressindo.
- Rismawan, Tedy. 2008. Aplikasi K-Means untuk Pengelompokan. *Jurnal Sistem dan Informatika*. Vol 4: 43 – 47
- Santosa, T. 2007. *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- Santoso, S. 2010. *Statistik Multivariat*. Jakarta: Elex Media Komputindo.