

ANALISIS PERBANDINGAN METODE SINGLE LINKAGE DAN K-MEANS CLUSTERING (STUDI KASUS : KASUS DEMAM BERDARAH DENGUE (DBD) DI PROVINSI JAWA TENGAH TAHUN 2015)

Denisha Intan P.¹⁾, Nurul Imani²⁾, Tiara Shafira³⁾, Linda Kurnia⁴⁾, Eka Rusnita⁵⁾, Edy Widodo⁶⁾

^{1,2)} Program Studi Statistika, Universitas Islam Indonesia Yogyakarta
Jl. Kaliurang KM 14.5, Umbulmartani, Ngemplak, Sleman, Yogyakarta 55584
Email : denisha665@gmail.com¹⁾, imaninurul47@gmail.com²⁾, tiarashafira60@gmail.com³⁾,
lindakurnia01@gmail.com⁴⁾, ekarusnita@gmail.com⁵⁾, 966110103@uii.ac.id⁶⁾

Abstrak

Penyakit Demam Berdarah Dengue (DBD) menjadi salah satu permasalahan kesehatan di Indonesia. Data dari Kementerian Kesehatan menyebutkan sebanyak 511 kabupaten/kota di Indonesia berpotensi menjadi tempat berkembangnya demam berdarah. Hal ini berarti tidak ada satu pun daerah Indonesia yang bebas terhadap endemisitas demam berdarah, termasuk provinsi Jawa Tengah. Pada tahun 2015, hampir semua kabupaten atau kota melaporkan adanya korban jiwa dikarenakan penyakit DBD. Tingginya kasus DBD di Jawa Tengah sehingga diperlukan adanya identifikasi penyebab masalah tersebut. Salah satu caranya yaitu dengan mengetahui apakah ada pengaruh antar kabupaten atau kota menggunakan Indeks Moran's dan LISA. Selain itu juga dilakukan pengelompokan setiap kabupaten atau kota mengenai kasus DBD berdasarkan jumlah kasus DBD, kepadatan penduduk, tingkat kesakitan, persentase rumah sehat, serta persentase rumah tangga ber-PHBS. Pada penelitian ini juga dilakukan perbandingan antara metode pengelompokan yang digunakan yaitu Single Linkage dan K-Means Clustering dengan melihat nilai simpangan baku terendah antar cluster untuk mendapatkan hasil pengelompokan terbaik. Hasil penelitian menunjukkan bahwa ada pengaruh antar kabupaten dan kota, serta didapatkan hasil pengelompokan sebanyak 4 cluster. Nilai simpangan baku cluster Single Linkage dan K-Means masing-masing adalah sebesar 0,679 dan 0,573. Berdasarkan nilai tersebut didapatkan bahwa metode K-Means Clustering sebagai metode terbaik dalam pengelompokan ini.

Kata kunci: DBD, Indeks Moran's, LISA, Single Linkage, K-Means.

1. Pendahuluan

Demam Berdarah Dengue (DBD) merupakan penyakit yang sampai saat ini menjadi salah satu masalah kesehatan di Indonesia. Penyakit ini pertama kali ditemukan di kota Surabaya pada tahun 1968 dengan sebanyak 58 orang terinfeksi dan 24 orang diantaranya meninggal dunia. Menurut data *World Health Organization* (WHO), Asia Pasifik menanggung 75%

dari beban *dengue* di dunia antara tahun 2004 dan 2010, sementara Indonesia dilaporkan sebagai negara ke-2 dengan kasus DBD terbesar diantara 30 negara wilayah endemis [1].

Data internal Pencegahan dan Pengendalian Penyakit (P2P) menyatakan penderita DBD pada tahun 2015 di 34 provinsi di Indonesia sebanyak 129.179 orang, dimana 1.240 diantaranya meninggal dunia. Data dari Kementerian Kesehatan menyebutkan sebanyak 511 kabupaten/kota di Indonesia berpotensi menjadi tempat berkembangnya demam berdarah [2]. Hal ini berarti tidak ada satu pun daerah Indonesia yang bebas terhadap endemisitas demam berdarah. Dari 500 kabupaten/kota yang berpotensi, hampir 90% diantaranya merupakan daerah endemik, termasuk provinsi Jawa Tengah.

Penyakit DBD masih merupakan permasalahan serius di provinsi Jawa Tengah, terbukti 35 kabupaten/kota sudah pernah terjangkit penyakit DBD. Angka kesakitan/*Indicate Rate* (IR) DBD provinsi Jawa Tengah pada tahun 2015 sebesar 47,9 per 100.000 penduduk [3]. Tingginya angka kesakitan DBD disebabkan karena adanya iklim tidak stabil dan curah hujan cukup banyak pada musim penghujan yang merupakan sarana perkembangbiakan nyamuk *Aedes Aegypti* yang cukup potensial. Selain itu juga didukung dengan tidak maksimalnya kegiatan Pemberantasan Sarang Nyamuk (PSN) di masyarakat sehingga menimbulkan Kejadian Luar Biasa (KLB) penyakit DBD di beberapa kabupaten/kota.

Pada penelitian yang dilakukan Faiz, dkk (2013) dalam Anisahtul (2016) sebelumnya menyatakan bahwa tingkat ketergantungan penyakit DBD di suatu daerah diperkirakan dipengaruhi oleh penyakit DBD di daerah lain yang berdekatan [4]. Pernyataan tersebut juga didukung oleh Hukum Geografi Pertama yang diungkapkan Tobler dalam Enselin (1993) yaitu "Semua hal berhubungan dengan hal lainnya, tetapi hal yang dekat lebih berhubungan dibandingkan dengan hal yang berjauhan". Oleh karena itu, jika suatu wilayah menjadi endemi penyakit DBD, maka diduga wilayah tersebut akan membuat wilayah yang berbatasan langsung dengannya menjadi endemi penyakit DBD yang baru. Berdasarkan penelitian Bakti (2012), analisis data spasial merupakan analisis yang berhubungan dengan pengaruh

lokasi. Pola hubungan antar lokasi yaitu wilayah satu dengan wilayah lainnya dapat diketahui dengan melihat nilai autokorelasi spasial. Beberapa pengujian dalam spasial autokorelasi spasial adalah menggunakan Indeks Moran (I) dan *Local Indicator of Spatial Autocorrelation* (LISA). Metode ini sangat penting untuk mendapatkan informasi mengenai pola penyebaran karakteristik suatu wilayah dan keterkaitan antar lokasi di dalamnya. Pada penelitian ini, dilakukan analisis spasial dengan mencari nilai autokorelasi spasial menggunakan *Indeks Moran* dan *LISA* untuk mengetahui persebaran kasus DBD di Jawa Tengah pada tahun 2015.

Selain itu, peneliti juga melakukan pengelompokan tingkat kerawanan DBD untuk menentukan daerah berdasarkan tingkat kerawanan DBD. Peneliti menggunakan metode *K-means* yang merupakan metode *clustering* yang paling sederhana dan umum. Hal ini dikarenakan *K-means* mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relatif cepat dan efisien. Peneliti juga menggunakan metode *hierarchical clustering* yaitu *single linkage clustering* yang merupakan metode terbaik pada penelitian yang dilakukan oleh Anisahtul (2016) sebelumnya. Kedua metode ini akan dibandingkan untuk melihat *cluster* mana yang memberikan hasil pengelompokan yang lebih baik. Proses pengelompokan ini nantinya diharapkan akan diketahui kemiripan atau kedekatan antar data sehingga dapat dikelompokkan ke dalam beberapa *cluster*, dimana antar anggota *cluster* memiliki tingkat kemiripan yang tinggi.

Tujuan Penelitian

Berdasarkan permasalahan diatas maka tujuan penelitian ini adalah :

1. Untuk mengetahui kelompok daerah berdasarkan tingkat kerawanan kasus DBD di Jawa Tengah tahun 2015.
2. Untuk mengetahui penyebaran kasus DBD di Jawa Tengah dengan menggunakan Indeks Moran dan LISA.
3. Untuk mengetahui metode terbaik dalam mengelompokkan tingkat kerawanan kasus DBD di Jawa Tengah tahun 2015.

Metode Penelitian

Sumber data dan Metode Analisis

Data yang digunakan dalam penelitian ini adalah data sekunder. Data tersebut diperoleh dari Skripsi Mahasiswa Statistika Universitas Islam Indonesia Tahun 2016 yang berjudul "*Analisis Regresi Spasial Faktor Pendorong Penyakit Demam Berdarah Dengue (DBD) Di Provinsi Jawa Tengah Tahun 2015*" dengan menggunakan variabel adalah Jumlah Kasus DBD, Kepadatan Penduduk, Indicate Ratio (IR) DBD, Persentase Rumah Sehat, dan Persentase Rumah Tangga Ber-PHBS. Data ini dianalisis menggunakan metode *cluster*. Analisis *cluster*

merupakan kelas teknik, dipergunakan untuk mengklasifikasi objek atau kasus (responden) ke dalam kelompok yang relatif homogen, yang disebut klaster (*clusters*). Objek/kasus dalam setiap kelompok cenderung mirip satu sama lain dan berbeda jauh (tidak sama) dengan objek dari *cluster* lainnya (Supranto, 2004) [5].

Secara umum, analisis *cluster* terbagi menjadi dua metode, yaitu metode Hierarki dan metode Non-Hierarki (Usman dan Sobari, 2013) [6]. Salah satu metode dalam metode Hierarki adalah metode *single linkage* dan dalam metode Non Hierarki adalah metode *K-means*. Metode pengelompokan *single linkage* dimulai dengan mengelompokkan dua objek yang mempunyai jarak terdekat. Misalkan objek X dan Y mempunyai jarak { d_{xy} } terdekat (Usman dan Sobari, 2013). Sedangkan pada metode *K-Means*, metode ini mengelompokkan objek dalam kelompok sedemikian rupa sehingga jarak tiap objek ke pusat kelompok minimum (Usman dan Sobari, 2013).

Algoritma *K-Means* merupakan algoritma klasterisasi yang mengelompokkan data berdasarkan titik pusat *cluster* (centroid) terdekat dengan data. Tujuan dari *K-Means* adalah pengelompokkan data dengan memaksimalkan kemiripan data dalam satu *cluster* dan meminimalkan kemiripan data antar *cluster*. Ukuran kemiripan yang digunakan dalam *cluster* adalah fungsi jarak. Sehingga pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data terhadap titik centroid.

Tahapan awal yang dilakukan pada proses klasterisasi dengan menggunakan algoritma *K-Means* adalah pembentukan titik awal centroid c_j . Pada umumnya pembentukan titik awal centroid dibangkitkan secara acak. Jumlah centroid c_j yang dibangkitkan sesuai dengan jumlah *cluster* yang ditentukan di awal. Setelah k centroid terbentuk kemudian dihitung jarak tiap data x_i dengan centroid ke- j sampai k , dinotasikan dengan $d(x_i, c_j)$. Terdapat beberapa ukuran jarak yang digunakan sebagai ukuran kemiripan suatu instance data, salah satunya adalah jarak Euclid. Perhitungan jarak Euclidean seperti pada persamaan berikut.

$$d(X_i, C_j) = \sqrt{\sum_{i=1}^N (X_i - C_j)^2}$$

Dengan :

$d(X_i, C_j)$: Jarak objek antara X_i dan C_j

N : Dimensi data

X_i : Koordinat dari objek X_i pada dimensi i

C_j : Koordinat dari objek Y_i pada dimensi i

Duran dan Odell (1974) menyatakan jika $d(X_i, C_j)$ semakin kecil, kesamaan antara dua unit pengamatan semakin dekat [7]. Syarat menggunakan jarak Euclid adalah jika semua fitur dalam dataset tidak saling

berkorelasi. Jika terdapat fitur yang berkorelasi maka menggunakan konsep jarak Mahalanobis.

Agusta (2007) menyatakan kelanjutan dari jarak tersebut dicari yang terdekat sehingga data akan mengelompok berdasarkan centroid yang paling dekat [8]. Tahap berikutnya adalah *update* titik centroid dengan menghitung rata-rata jarak seluruh data terhadap centroid. Selanjutnya akan kembali lagi ke proses awal. Iterasi ini akan diulangi terus sampai didapatkan centroid yang konstan artinya titik centroid sudah tidak berubah lagi. Atau iterasi dihentikan berdasarkan jumlah iterasi maksimal yang ditentukan.

Pada Analisis Cluster terdapat dua asumsi dalam analisis *cluster*, yaitu :

- a. Populasi atau sampel representative
 Data yang digunakan dalam analisis *cluster* adalah data populasi. Data sampel dapat digunakan namun harus representative yang artinya seberapa besar sampel harus mewakili populasi. Untuk mengetahui hal itu dilihat dari nilai *Kaisen Meyer Olkin* (KMO). KMO adalah indeks perbandingan nilai koefisien korelasi terhadap korelasi parsial.

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \dots\dots (1)$$

$$a_{ij}^2 = \frac{-r_{ij}}{\sqrt{r_{ij}r_{ij}}} \dots\dots(2)$$

dengan,

- p = banyaknya variabel
- r_{ij} = koefisien korelasi antara variabel i dan j
- a_{ij} = koefisien korelasi parsial antara variabel i dan j

Apabila nilai $KMO < 0.5$, maka sampel tidak mewakili populasi, sedangkan jika nilai $KMO > 0.5$, maka sampel mewakili populasi sehingga layak untuk dilakukan analisis *cluster*.

- b. Tidak ada multikolinearitas
 Multikolinearitas adalah adanya hubungan linear atau korelasi yang tinggi antar variabel. Untuk mengetahui adanya multikolinearitas dapat dilihat dari nilai-nilai korelasi pada matriks korelasi. Dikatakan multikolinearitas jika nilai korelasi antar variabel $> 0,70$ dan dapat diatasi menggunakan *Principal Component Analysis* (PCA). Kelebihan dari PCA yaitu dapat mengatasi multikolinearitas secara bersih. Artinya, *Principal Component* (PC) hasil dari PCA bebas dari multikolinearitas.

Analisis Data Spasial

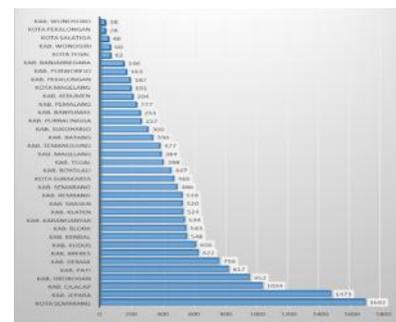
Menurut Pfeiffer dkk (2008) dalam Rahmawati (2013) [9], analisis data spasial pada umumnya dilakukan untuk setiap wilayah atau lokasi. Hal ini dikarenakan setiap wilayah memiliki karakteristik sendiri sehingga dalam

pendekatan analisisnya menggunakan eksplorasi data yang disajikan dalam bentuk peta tematik.

Moran's Scatter plot adalah alat yang digunakan untuk melihat hubungan antara nilai pengamatan yang terstandarisasi dengan rata-rata tetangga yang sudah terstandarisasi. Jika digabungkan garis regresi maka dapat digunakan untuk mengetahui derajat kecocokan dan mengidentifikasi adanya *outlier*.

Menurut Lee dan Wong dalam Sutikno (2011) dalam Anisahtul (2016), untuk mengidentifikasi koefisien autokorelasi dapat menggunakan Indeks Moran pada LISA. LISA mengidentifikasi bagaimana hubungan antara suatu lokasi pengamatan terhadap lokasi pengamatan lainnya.

2. Pembahasan



Gambar 1. Grafik Jumlah Kasus DBD Tahun 2015

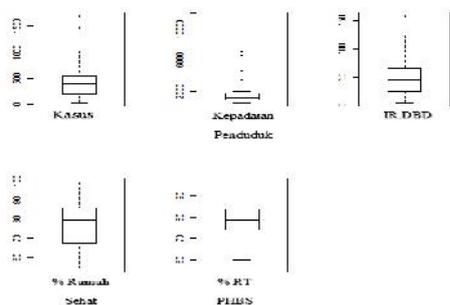
Gambar 1 menunjukkan bahwa jumlah kasus DBD tertinggi di Provinsi Jawa Tengah terjadi di Kota Semarang yaitu sebanyak 1.692 kasus. Posisi kedua dan ketiga diduduki oleh Kabupaten Jepara dan Kabupaten Cilacap secara berturut-turut sebanyak 1.473 dan 1.054 kasus. Pada ketiga daerah ini hal tersebut diduga disebabkan oleh perubahan cuaca yang tidak menentu yang menyebabkan jentik nyamuk *Aedes aegypti* bisa berkembang dengan cepat.

Untuk melakukan analisis *cluster* maka kedua asumsi pada analisis ini harus terpenuhi dan melakukan pendeteksian data *outlier*.

Tabel 1 Nilai *Variance Inflation Multikolinearitas* untuk Setiap Variabel

Variabel	Nilai VIF	Keputusan
Jumlah Kasus DBD	1,944	Tidak multikolinearitas
Kepadatan Penduduk	1,772	Tidak multikolinearitas
Indicate Rate	2,184	Tidak multikolinearitas
Persentase Rumah Sehat	1,451	Tidak multikolinearitas
Persentase Rumah Tangga Ber-PHBS	1,627	Tidak multikolinearitas

Berdasarkan tabel 1 diperoleh nilai VIF (*Variance Inflation Multikolinearitas*) < 10 yang artinya bahwa tidak mengalami multikolinearitas antar variabel.



Gambar 2. Boxplot Variabel

Berdasarkan gambar 2, didapatkan bahwa pada variabel Kasus, Kepadatan Penduduk, IR DBD, dan Persentase Rumah Tangga-berPHBS mengandung data outlier. Selanjutnya peneliti memisahkan data outlier tersebut membentuk 1 cluster dan membentuk cluster lain berdasarkan data-data yang tidak mengandung outlier sebanyak 3 cluster, sehingga total cluster yang dibentuk adalah sebanyak 4 cluster baik untuk metode *Single Linkage* maupun *K-Means*. Berdasarkan data hasil output cluster membership diperoleh anggota dari masing-masing cluster dapat dilihat pada table berikut.

Tabel 2. Anggota Masing-Masing Cluster Metode *Single Linkage*

C	Anggota Cluster (Kabupaten/Kota)
1	Kab. Cilacap, Kab. Banyumas, Kab. Purbalingga, Kab. Kebumen, Kab. Purworejo, Kab. Wonosobo, Kab. Magelang, Kab. Boyolali, Kab. Klaten, Kab. Sukoharjo, Kab. Wonogiri, Kab. Karanganyar, Kab. Sragen, Kab. Blora, Kab. Rembang, Kab. Pati, Kab. Kudus, Kab. Demak, Kab. Semarang, Kab. Temanggung, Kab. Kendal, Kab. Batang, Kab. Pekalongan, Kab. Pemalang, Kab. Tegal.
2	Kab. Banjarnegara
3	Kab. Grobogan
4	Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, Kota Tegal, Kab. Jepara, Kab. Brebes

Berdasarkan nilai rata-rata untuk masing-masing variabel pada masing-masing cluster pada metode ini dapat diinterpretasikan hasil sebagai berikut :

1) *Cluster 1* : Pada cluster 1 terdiri dari 25 kabupaten/kota. Cluster 1 merupakan kelompok yang memiliki tingkat kerawanan DBD terendah nomor 2. Karakteristik ini didapat dari hasil perhitungan yang menunjukkan bahwa rata-rata jumlah kasus DBD cluster 1 lebih rendah dari cluster 3 dan 4 tetapi lebih tinggi daripada cluster 2. Hal ini bisa disebabkan karena baiknya aspek lain yang ditunjukkan dengan

rata-rata persentase rumah sehat yang baik dan persentase rumah tangga ber-PHBS yang berada diatas rata-rata rumah tangga ber-PHBS cluster lain. Dari hasil tersebut, dapat di dikatakan bahwa daerah dengan persentase rumah sehat dan rumah tangga ber-PHBS tinggi dapat meminimalkan jumlah kasus DBD yang terjadi. Sehingga dalam pengendalian kasus DBD, memiliki rumah sehat dan rumah tangga ber-PHBS dapat diterapkan untuk meminimumkan terjadinya kasus DBD.

- 2) *Cluster 2* : Pada cluster 2 ini hanya terdiri dari 1 kabupaten, yaitu Kabupaten Banjarnegara. Cluster 2 memiliki tingkat kerawanan kasus DBD paling rendah. Hal ini juga ditunjukkan dengan karakteristik jumlah kasus penyakit DBD paling rendah daripada cluster lainnya. Jika dilihat dari kepadatan penduduk, cluster 2 memiliki kepadatan penduduk yang rendah.
- 3) *Cluster 3* : Pada cluster 3 hanya terdiri dari 1 kabupaten yaitu Kabupaten Grobogan dengan tingkat kerawanan DBD sedang. Jika dilihat dari jumlah kasus DBD, cluster ini memiliki jumlah yang paling tinggi dari cluster lainnya, tetapi kondisi rumah sehat cluster ini tergolong tinggi dan cluster ini juga memiliki persentase rumah tangga ber-PHBS paling tinggi. Hal ini mempertegas bahwa meskipun suatu daerah memiliki jumlah kasus DBD yang tinggi jika diimbangi dengan kondisi rumah sehat dan persentase rumah tangga ber-PHBS yang tinggi maka dapat meminimalisasikan tingkat kerawanan DBD daerah tersebut.
- 4) *Cluster 4* : Pada cluster 4 terdiri dari 8 kota/kabupaten yang memiliki karakteristik berbeda dari daerah lainnya karena pada setiap variabel pada anggota cluster ini memiliki karakteristik yang jauh dari rata-rata sehingga anggota cluster 4 memerlukan penanganan tersendiri. Pada cluster 4 memiliki nilai kerawanan DBD paling tinggi dibandingkan ketiga cluster lainnya. Karakteristik ini dilihat dari rata-rata kasus DBD cluster 4 tertinggi nomor 2 setelah cluster 3 ditambah dengan jumlah kepadatan penduduk paling tinggi diantara cluster lainnya. Ini berarti bahwa daerah dengan kepadatan penduduk tinggi dapat dikatakan sebagai daerah dengan tingkat kerawanan DBD tinggi.

Tabel 3. Hasil Pengelompokkan *K-Means*

C	Anggota
1	Kab. Cilacap, Kab. Rembang, Kab. Pati, Kab. Kudus, Kab. Semarang, Kab. Kendal
2	Kab. Banyumas, Kab. Banjarnegara, Kab. Kebumen, Kab. Wonosobo, Kab. Magelang, Kab. Boyolali, Kab. Wonogiri, Kab. Batang, Kab. Pekalongan, Kab. Pemalang, Kab. Tegal
3	Kab. Purbalingga, Kab. Purworejo, Kab. Klaten, Kab. Sukoharjo, Kab. Karanganyar, Kab. Sragen, Kab.

	Grobogan, Kab. Blora, Kab. Demak, Kab. Temanggung
4	Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, Kota Tegal, Kab. Jepara, Kab. Brebes

Pemilihan Metode Terbaik dengan Simpangan Baku

Tabel 4. Perbandingan Nilai Rasio Simpangan Baku

Metode	Nilai Simpangan Baku
<i>Single Linkage</i>	0,679
<i>Kmeans</i>	0,573

Berdasarkan perbandingan nilai rasio simpangan baku pada tabel 4 menunjukkan bahwa metode terbaik dalam mengelompokkan daerah berdasarkan tingkat kerawanan DBD di Jawa Tengah dengan menggunakan 4 cluster adalah metode *K-Means* karena memiliki nilai rasio lebih kecil dibanding metode *Single Linkage Clustering*.

Indeks Moran dan LISA

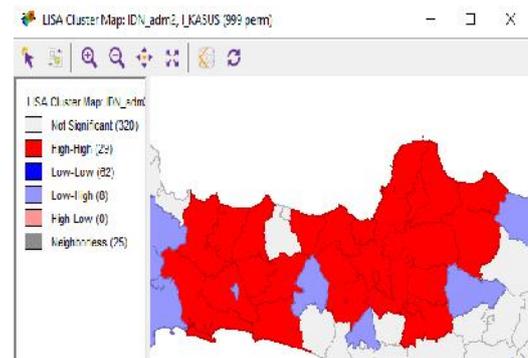
Indeks Moran merupakan salah satu teknik spasial yang digunakan untuk melihat adanya autokorelasi spasial antar lokasi pengamatan (Wuryandari dkk, 2014). Adapun perhitungan untuk Indeks Moran menggunakan *software* Geoda dengan *output* berupa nilai Indeks Moran, yaitu sebagai berikut:

Tabel 5. Hasil Pengujian Indeks Moran

Variabel	Indeks Moran
Kasus DBD	0,624
Kepadatan Penduduk	0,246
IR DBD	0,602
Rumah Sehat	0,8
Rumah Tangga Ber-PHBS	0,802

$$I_0 = -\frac{1}{n-1} = -\frac{1}{35-1} = -0,029411$$

Berdasarkan 5 dan nilai I_0 didapatkan hasil bahwa seluruh variabel memiliki nilai Indeks Moran lebih besar dari n_i ; I_0 sehingga dapat diambil kesimpulan bahwa seluruh variabel memiliki autokorelasi spasial positif. Artinya variabel kasus DBD, kepadatan penduduk, IR DBD, rumah sehat dan rumah tangga ber-PHBS memiliki pola pengelompokan dan memiliki karakteristik pada lokasi yang berdekatan.



Gambar 6 LISA Kasus DBD

Berdasarkan hasil pengujian LISA pada gambar 6 dapat dilihat bahwa kasus DBD cenderung berada di seluruh wilayah Jawa Tengah hal ini dibuktikan dengan wilayah yang berada pada area *hot spot* ditunjukkan dengan wilayah berwarna merah, merupakan wilayah yang tinggi dengan ketetaanggan tinggi dimana jumlah untuk wilayah ini sebanyak 29 kabupaten. Kabupaten yang masuk dalam area *hot spot* ini perlu mendapatkan perhatian khusus dari pemerintah karena persentase kasus DBD yang terjadi tinggi dan juga dikelilingi oleh kabupaten lain yang memiliki jumlah kasus DBD yang tinggi pula. Sedangkan untuk wilayah berwarna biru muda merupakan area *cold spot* yaitu wilayah yang rendah dengan ketetangaan tinggi dengan jumlah sebanyak 8 kabupaten. Kasus DBD pada area *cold spot* ini juga memerlukan perhatian dari pemerintah karena kasus DBD yang terjadi di wilayah ini didukung oleh jumlah kasus DBD yang rendah tetapi dipengaruhi oleh kabupaten yang memiliki jumlah kasus DBD yang tinggi.

3. Kesimpulan

- Berdasarkan hasil *clustering* menggunakan metode *single linkage*, diperoleh pengelompokan tingkat kerawanan kasus DBD di Jawa Tengah tahun 2015 dengan jumlah anggota pada *cluster* 1 sebanyak 25 kabupaten/kota, *cluster* 2 dua dan 3 masing-masing sebanyak satu daerah, sedangkan *cluster* 4 terdiri dari 8 daerah yang memiliki karakteristik berbeda dari *cluster* lainnya dikarenakan adanya data *outlier*. Sementara dengan menggunakan metode *K-Means clustering*, diperoleh jumlah anggota untuk setiap *cluster* adalah 6 daerah pada *cluster* 1, kemudian 11 daerah pada *cluster* 2, dan 10 daerah pada *cluster* 3 dan 8 daerah pada *cluster* 4.
- Hasil dari nilai *indeks Moran's* sebesar 0,624 itu artinya bahwa autokorelasi yang dihasilkan adalah autokorelasi spasial positif dimana antar kabupaten satu dengan kabupaten lain berkelompok. Sedangkan dari hasil pengujian LISA banyak wilayah cenderung berada di area *hot spot* sebanyak 29 kabupaten dan nilai *cold spot* sebanyak 8 kabupaten.
- Dalam penelitian ini diperoleh hasil perbandingan dua metode *cluster single linkage* dan *K-Means clustering*, diperoleh bahwa berdasarkan

perbandingan nilai rasio simpangan baku pada metode terbaik dalam mengelompokkan daerah berdasarkan tingkat kerawanan DBD di Jawa Tengah dengan menggunakan 4 *cluster* adalah metode *K-Means clustering* karena nilai rasio yang dimiliki metode *K-Means clustering* lebih kecil dari metode *single linkage* yaitu sebesar 0.573

Daftar Pustaka

- [1] CNN Indonesia, "Indonesia Peringkat Dua Negara Endemis Demam Berdarah," diakses dari <https://www.cnnindonesia.com/gaya-hidup/20160616170332-255-138672/indonesia-peringkat-dua-negara-endemis-demam-berdarah>, pada tanggal 28 September 2017.
- [2] CNN Indonesia, "Kemenkes: Indonesia Belum Bebas Demam Berdarah," diakses dari <https://www.cnnindonesia.com/gaya-hidup/20160112160559-255-103778/kemenkes-indonesia-belum-bebas-demam-berdarah>, pada tanggal 30 September 2017.
- [3] Dinas Kesehatan Provinsi Jawa Tengah, "Profil Kesehatan Jawa Tengah Tahun 2015," diakses dari <http://www.dinkesjatengprov.go.id/v2015/dokumen/profil2015/ind-ex.html>, pada tanggal 29 September 2017.
- [4] Fithriyyah, A, "Analisis *Cluster* Spasial Tingkat Kerawanan Demam Berdarah Dengue (DBD) di Provinsi Jawa Tengah Tahun 2015," Skripsi Program Studi Statistika FMIPA Universitas Islam Indonesia.
- [5] Supranto, J, "Analisis Multivariat, Arti dan Interpretasi", Jakarta: PT Rineka Cipta, 2004.
- [6] Usman, H., dan Nurdin S, "Aplikasi Teknik Multivariate untuk Riset Pemasaran", Jakarta: PT RajaGrafindo Persada, 2013.
- [7] Duran, Benjamin S. dan Odell, Patrick L., Cluster Analysis, 1974.
- [8] Agusta, Yudi., "K-Means Penerapan, Permasalahan dan Metode Terkait", Jurnal Sistem dan Informatika 3.47-60, 2007
- [9] Rahmawati, L, "Analisis Kelompok dengan Menggunakan Metode Hierarki untuk Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kesehatan," Skripsi Jurusan Matematika FMIPA Universitas Negeri Malang.

Biodata Penulis

Denisha Intan P., Nurul Imani, Tiara Shafira, Linda Kurnia, Eka Rusnita, Mahasiswa Statistika Universitas Islam Indonesia.

Dr. Edy Widodo, S.Si., M.Si., memperoleh gelar Sarjana Sains (S.Si), Jurusan Statistika Institut Sepuluh November, lulus tahun 1996. Memperoleh gelar Magister Sains (M.Si) Program Pasca Sarjana Magister Matematika Universitas Gajah Mada, lulus tahun 2001. Memperoleh gelar Doktor (Dr) Program Doktoral Matematika Universitas Gajah Mada, lulus tahun 2015. Saat ini menjadi Dosen di Statistika Universitas Islam Indonesia.