

Received: 27 January 2023

Revised: 14 June 2023

Accepted: 28 June 2023

Published: 30 June 2023

Analisis Faktor-Faktor yang Menjelaskan Kasus AIDS Provinsi Jawa Timur Menggunakan Model *Geographically Weighted Logistic Regression* (GWLR)

Natasha Latifatu Soliha^{1, a)}, Dian Lestari^{1, b)}, Yekti Widyaningsih^{1, c)}

¹Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indonesia

Email: ^{a)}natasha.latifatu@sci.ui.ac.id, ^{b)}dian.lestari@sci.ui.ac.id, ^{c)}yekti@sci.ui.ac.id

Abstract

AIDS is the most chronic phase of HIV infection which can weaken the immune system. In 2020, East Java Province is a province which has the most HIV infections and in the third place for the highest total number of AIDS cases in Indonesia. The purpose of this research is to build a model using Geographically Weighted Logistic Regression (GWLR), and to work out the grouping results of regencies/cities using K-means Clustering Analysis. The variables used in this research are Gini Ratio, L Index of Per Capita Expenditure, Gender Ratio, Dependency Ratio, Gender Development Index, and The Number of Pos Pelayanan KB Desa. The proportion levels of AIDS cases are categorized into 2 categories based on cut-point which has been specified, which 0 as the category of low level with the proportion of AIDS cases is less than 0.0006 and 1 as the category of high level with the proportion of AIDS cases is more than or equal to 0.0006. Parameter estimation for GWLR is using Maximum Likelihood Estimation (MLE) method with Fixed Gaussian as weighted kernel function and optimum bandwidth is determined using Akaike's Information Criterion Corrected (AICc). Z-Score of the most suitable model will be grouped using K-means Clustering Analysis, with Z-score is parameter estimator divided by standard error. Grouping results indicates cluster 1 members tend to be regencies/cities that have gender ratio and dependency ratio as significant variables, meanwhile cluster 2 members tend to be regencies/cities that have only dependency ratio as significant variable.

Keywords: *Fixed Gaussian; K-means Clustering; Maximum Likelihood; Spatial*

Abstrak

AIDS merupakan fase infeksi dari virus HIV paling kronis yang dapat melemahkan kekebalan tubuh. Pada tahun 2020, Provinsi Jawa Timur merupakan provinsi dengan jumlah paling tinggi pada infeksi HIV dan urutan ketiga terbanyak untuk total kasus AIDS di Indonesia. Tujuan penelitian ini untuk memodelkan data tingkat proporsi kasus AIDS Provinsi Jawa Timur menggunakan *Geographically Weighted Logistic Regression* (GWLR) dan melihat hasil pengelompokan kabupaten/kota menggunakan analisis kluster *k-means*. Variabel yang digunakan

untuk penelitian ini adalah Gini Rasio, Indeks L Pengeluaran Per Kapita, Rasio Jenis Kelamin, Rasio Ketergantungan, Indeks Pembangunan Gender (IPG), dan Jumlah Pos Pelayanan KB Desa. Tingkat proporsi kasus AIDS dikategorikan menjadi 2 kategori berdasarkan *cut-point* yang telah ditentukan, dengan kategori 0 sebagai tingkat rendah dengan proporsi kasus AIDS kurang dari 0,0006 dan kategori 1 sebagai tingkat tinggi dengan proporsi kasus AIDS lebih dari atau sama dengan 0,0006. Penaksiran parameter untuk model GWLR menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan fungsi pembobot kernel *Fixed Gaussian* dan *bandwidth* optimum ditentukan menggunakan *Akaike's Information Criterion corrected* (AICc). Nilai *Z* hitung dari parameter model yang paling sesuai akan dikelompokkan menggunakan analisis kluster *k-means*, dengan *Z* hitung adalah nilai estimasi parameter dibagi dengan standar error. Hasil pengelompokan menunjukkan bahwa anggota kluster 1 memiliki kecenderungan merupakan kabupaten/kota yang memiliki variabel signifikan yaitu rasio jenis kelamin dan rasio ketergantungan, sementara anggota kluster 2 memiliki kecenderungan merupakan kabupaten/kota yang memiliki variabel signifikan rasio ketergantungan.

Kata-kata kunci: *fixed gaussian*; kluster *k-means*; *maximum likelihood*; spasial

PENDAHULUAN

HIV (*Human Immunodeficiency Virus*) adalah sejenis virus yang menyerang sistem imun pada tubuh. Saat infeksi HIV semakin menurunkan sistem kekebalan tubuh, virus tersebut menumbuhkan tanda dan gejala lain, seperti diare, demam, batuk, menurunnya berat badan secara drastis, dan juga adanya peningkatan kelenjar getah bening. HIV juga dapat mengakibatkan terjadinya penyakit AIDS (*Acquired Immune Deficiency Syndrome*), dimana AIDS adalah fase terakhir atau fase infeksi paling kronis dari virus HIV (*National Institute of Health*, 2021).

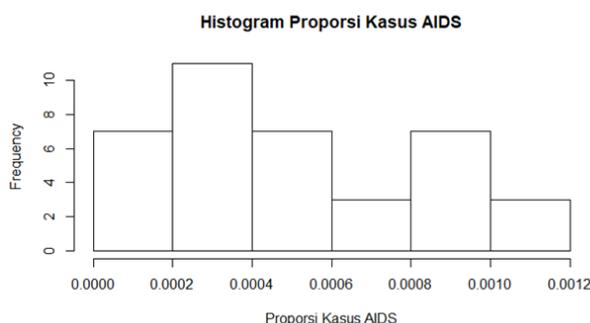
Jumlah kasus HIV/AIDS di Indonesia cenderung turun naik. Pada tahun 2019, kasus HIV di Indonesia sangat tinggi, yaitu sebanyak 50.282 kasus. Jumlah infeksi HIV yang paling tinggi berada di Provinsi Jawa Timur. Sementara itu, Provinsi Jawa Timur berada di deretan ketiga terbanyak untuk jumlah kasus AIDS yang dilaporkan dengan jumlah 958 kasus atau sekitar 13,62% dari total kasus AIDS yang terjadi pada tahun 2019 (Pusdatin Kemenkes RI, 2020). Pada penelitian sebelumnya yang dilakukan oleh Ratnasari dan Purnadi (2013) dengan judul "Pemodelan Faktor yang Mempengaruhi Jumlah HIV dan AIDS Provinsi Jawa Timur Menggunakan Regresi Poisson Bivariat" didapat faktor-faktor yang berpengaruh terhadap kasus HIV dan AIDS di kabupaten/kota di Jawa Timur adalah persentase pengguna kondom, persentase penduduk kelompok umur 25-29 tahun, persentase daerah berstatus desa, persentase penduduk tamat SMA, dan persentase penduduk miskin di tiap kabupaten/kota.

Pada penelitian ini, faktor-faktor yang diduga berpengaruh terhadap kasus AIDS Jawa Timur antara lain, gini rasio, pengeluaran per kapita, rasio jenis kelamin, rasio ketergantungan, indeks pembangunan gender, dan jumlah pelayanan pos KB desa di setiap kabupaten/kota. Setiap kabupaten/kota di Jawa Timur memiliki proporsi jumlah kasus AIDS terhadap jumlah penduduk yang berbeda-beda yang dibagi menjadi dua tingkatan, yaitu tinggi dan rendah. Model yang tepat untuk menganalisis variabel respon dengan 2 kategori adalah regresi logistik biner (Agresti, 2012). Analisis regresi logistik biner atau biasa disebut juga dengan *logit model* diperlukan untuk menganalisis hubungan satu atau beberapa variabel prediktor dengan sebuah variabel respon yang bersifat biner atau dikotomi (Hosmer, Lemeshow, & Sturdivant, 2013). Selain itu, antar kabupaten/kota terkadang memiliki kondisi yang berbeda satu dengan lainnya. Terdapatnya pengaruh aspek geografis atau spasial pada suatu lokasi mengakibatkan kemungkinan adanya heterogenitas spasial. Heterogenitas spasial ialah keadaan dimana terdapat keberagaman hubungan secara kewilayahan. Metode yang dapat digunakan pada data spasial yang memiliki heterogenitas spasial adalah *Geographically Weighted Regression* (GWR), dimana *Geographically Weighted Logistic Regression* (GWLR) adalah salah satu pengembangan dari metode GWR (Fotheringham, dkk., 2002). Sehingga, analisis data pada penelitian ini dilakukan dengan menggunakan model *Weighted Logistic Regression* (GWLR) untuk menganalisis faktor-faktor yang menjelaskan tingkat proporsi kasus AIDS di kabupaten/kota di Provinsi Jawa Timur pada tahun 2020.

METODOLOGI

Bahan dan Data

Penelitian ini menggunakan data sekunder yang bersumber pada situs web resmi Badan Pusat Statistik (BPS) pada tahun 2020 dengan wilayah penelitian Provinsi Jawa Timur yang terdiri dari 38 kabupaten/kota. Tingkat proporsi kasus AIDS (Y) ditentukan berdasarkan histogram proporsi kasus AIDS per kabupaten/kota. Histogram untuk proporsi kasus AIDS dapat dilihat pada GAMBAR 1.



GAMBAR 1. Histogram Proporsi Kasus AIDS

Berdasarkan histogram pada GAMBAR 1, proporsi sebesar 0,0006 ditetapkan sebagai *cut point*. Sehingga, variabel respon (Y) pada penelitian ini adalah tingkat proporsi kasus AIDS dengan kategori 0 adalah kabupaten/kota dengan tingkat proporsi kasus AIDS yang rendah dan kategori 1 adalah kabupaten/kota dengan tingkat proporsi kasus AIDS tinggi. Variabel-variabel yang akan digunakan dapat dilihat pada TABEL 1.

TABEL 1. Variabel yang digunakan

| Nama Variabel | Keterangan |
|--|---|
| Tingkat Proporsi Kasus AIDS (Y) | Kategori tinggi atau rendahnya kasus AIDS. 1: Tinggi (proporsi kasus AIDS $\geq 0,0006$) 0: Rendah (proporsi kasus AIDS $< 0,0006$) |
| Gini Rasio (GR) | Alat ukur yang menunjukkan tingkat ketimpangan pendapatan penduduk secara menyeluruh (BPS, 2022). |
| Indeks L Pengeluaran Per Kapita (IL.PKP) | Ukuran kesenjangan pengeluaran per kapita dalam distribusi rumah tangga (BPS, 2022). |
| Rasio Jenis Kelamin (SR) | Perbandingan antara jumlah penduduk pria dan jumlah penduduk wanita pada suatu wilayah dan waktu tertentu (BPS, 2022). |
| Rasio Ketergantungan (RK) | Perbandingan jumlah penduduk yang belum produktif dan sudah tidak produktif lagi dengan jumlah penduduk yang masih produktif (BPS, 2022). |
| Indeks Pembangunan Gender (IPG) | Rasio capaian antara IPM perempuan dan laki-laki (BPS, 2022). |
| Jumlah Pos Pelayanan KB Desa (PPKB) | Jumlah sarana pembinaan dan pelayanan KB di desa. |

Metode Penelitian

1. Uji Multikolinieritas

Terjadinya multikolinieritas dapat dideteksi dengan mencari nilai *Variance Inflation Factor* (VIF). Rumus perhitungan VIF adalah sebagai berikut:

$$VIF(k) = \frac{1}{1 - R_k^2}, k = 1, 2, \dots, p \tag{1}$$

dengan R_k^2 adalah koefisien determinasi dari model regresi variabel prediktor ke- k , dimana $0 \leq R_k^2 \leq 1$, sehingga $VIF(k) \geq 1, \forall k$. Adanya multikolinieritas pada model regresi ketika nilai VIF lebih besar dari 10 (Medenhall dan Sincich, 2012).

2. Uji Heterogenitas Spasial

Heterogenitas dapat dideteksi menggunakan uji *Breusch-Pagan* (BP). Hipotesis pengujian heterogenitas spasial yaitu,

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2 \text{ (Tidak terdapat heterogenitas spasial)}$$

$$H_1: \text{Setidaknya terdapat satu } \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, n \text{ (Terdapat heterogenitas spasial)}$$

Statistik uji untuk uji heterogenitas spasial menggunakan uji Breusch-Pagan adalah sebagai berikut:

$$BP = \frac{1}{2} f^T Z(Z^T Z)^{-1} Z^T f \tag{3}$$

dengan elemen vektor f adalah sebagai berikut:

$$f_i = \left(\frac{e_i^2}{\sigma^2} - 1 \right) \tag{4}$$

dimana e_i^2 adalah *least square residual* untuk observasi ke- i , σ^2 adalah variansi residual, dan Z adalah matriks yang berisi vektor berdistribusi normal standar baku untuk setiap observasi dengan ukuran $n \times (p + 1)$. Daerah penolakan H_0 untuk uji *Breusch-Pagan* adalah ketika $BP > \chi_p^2$ atau $p - \text{value} < \alpha$, dengan χ_p^2 adalah nilai kritis dari tabel distribusi *chi-square* dengan derajat bebas p .

3. Geographically Weighted Logistic Regression (GWLR)

Model *Geographically Weighted Logistic Regression* (GWLR) dapat dinyatakan sebagai berikut:

$$\pi(x_i) = \frac{\exp(x_i^T \beta(u_i, v_i))}{1 + \exp(x_i^T \beta(u_i, v_i))} \tag{5}$$

dengan $\beta(u_i, v_i) = [\beta_0(u_i, v_i) \quad \beta_1(u_i, v_i) \quad \beta_2(u_i, v_i) \quad \dots \quad \beta_p(u_i, v_i)]^T$ dan $x_i^T = [1 \quad x_{i1} \quad x_{i2} \quad \dots \quad x_{ip}]$.

4. Uji Kesesuaian Model

Hipotesis pengujian untuk menguji kesamaan model sebagai berikut:

$$H_0: \beta_k(u_i, v_i) = \beta_k; i = 1, 2, \dots, n; k = 1, 2, \dots, p \text{ (Tidak terdapat perbedaan yang signifikan antara kedua model)}$$

$$H_1: \text{minimal terdapat satu } \beta_k(u_i, v_i) \neq \beta_k \text{ (Terdapat perbedaan yang signifikan antara kedua model)}$$

Statistik uji kesamaan model menggunakan perbandingan nilai devians masing-masing model, yaitu:

$$F_{hit} = \frac{D(\hat{\beta})/df_1}{D(\hat{\beta}^*)/df_2} \tag{6}$$

dengan $D(\hat{\beta})$ adalah nilai devians dari model regresi logistik biner global, $D(\hat{\beta}^*)$ adalah nilai devians dari model GWLR, $df_1 = p$, dan $df_2 = np$ (Caraka dan Yasin, 2017). Kriteria uji dari pengujian

kesesuaian model di atas adalah tolak H_0 jika $F_{hit} > F_{\alpha;df_1,df_2}$ dimana $F_{\alpha;df_1,df_2}$ adalah nilai kritis yang didapat dari tabel F dengan α adalah taraf signifikansi, derajat bebas df_1 adalah p dan df_2 adalah np , atau jika $p - value < \alpha$ pada taraf signifikansi α .

5. Uji Signifikansi Parameter

a. Uji Serentak

Hipotesis pengujian signifikansi parameter model secara serentak adalah sebagai berikut:
 $H_0: \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_p(u_i, v_i) = 0$ (Tidak terdapat pengaruh variabel independen terhadap variabel dependen)

H_1 : Minimal terdapat satu $\beta_k(u_i, v_i) \neq 0, k = 1, 2, \dots, p$ (Terdapat pengaruh variabel independen terhadap variabel dependen)

Statistik uji untuk pengujian parameter secara serentak adalah:

$$G_2 = -2 \ln \frac{L(\hat{\omega}_{GWLR})}{L(\hat{\Omega}_{GWLR})} \tag{7}$$

dengan $L(\hat{\omega}_{GWLR})$ adalah nilai *log-likelihood* tanpa variabel prediktor dan $L(\hat{\Omega}_{GWLR})$ adalah nilai *log-likelihood* dengan variabel prediktor untuk model GWLR. Aturan keputusan yang digunakan adalah H_0 ditolak jika $G_2 > \chi_{\alpha,v}^2$, dimana α adalah taraf signifikansi dan v adalah derajat bebas.

b. Uji Parsial

Hipotesis untuk pengujian parameter model secara parsial adalah sebagai berikut:

$H_0: \beta_k(u_i, v_i) = 0$

$H_1: \beta_k(u_i, v_i) \neq 0; k = 1, 2, \dots, p$

Statistik uji untuk pengujian parameter secara parsial adalah:

$$Z_{hit} = \frac{\hat{\beta}_k(u_i, v_i)}{se(\hat{\beta}_k(u_i, v_i))} \tag{8}$$

dengan $\hat{\beta}_k(u_i, v_i)$ adalah penaksir parameter ke- k untuk model GWLR dan $se(\hat{\beta}_k(u_i, v_i))$ adalah nilai *standar error* untuk parameter $\hat{\beta}_k(u_i, v_i)$. Aturan keputusan yang digunakan adalah H_0 ditolak jika $|Z_{hit}| > Z_{\frac{\alpha}{2}}$, dimana Z adalah nilai kritis yang didapat dari tabel distribusi normal.

6. Analisis Kluster *K-Means*

K-means merupakan salah satu algoritma *clustering* non-hirarki yang pengelompokkannya berdasarkan kesamaan karakteristik. Kluster yang terbentuk diwakili oleh *centroid* yang merupakan rata-rata dari titik-titik pada kluster tersebut (Wu, 2012). Algoritma *k-means* adalah metode *clustering* berdasarkan jarak yang terdekat dengan titik pusat yang ditentukan. Jarak yang sering digunakan adalah jarak *euclidean* yang dapat diperoleh dengan persamaan sebagai berikut:

$$d_{ij} = \sqrt{\sum_k^p (x_{ik} - x_{jk})^2} = \|\mathbf{x}_i - \mathbf{x}_j\|^2 \tag{7}$$

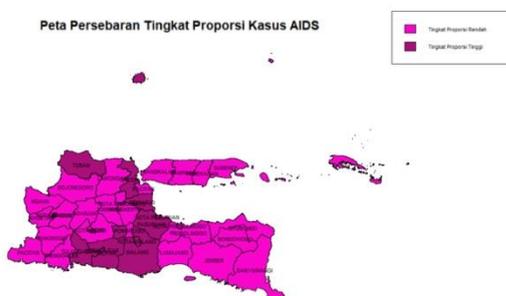
dengan d_{ij} adalah jarak *euclidean* antara objek i dan j , x_{ij} adalah nilai objek i untuk variabel ke- j , x_{jk} adalah nilai objek j untuk variabel ke- k , dan p adalah banyaknya variabel yang diamati. Tahapan metode analisis kluster *k-means* adalah sebagai berikut:

1. Menentukan jumlah *cluster* atau nilai k yang diinginkan
2. Menentukan pusat atau *centroid*

3. Menghitung jarak antar objek ke masing-masing pusat atau *centroid*
4. Mengelompokkan objek berdasarkan jarak terdekat dengan pusat *cluster*
5. Kembali menentukan *centroid* sampai semua objek membentuk kelompok atau *cluster*
(Abonyi dan Feil, 2007)

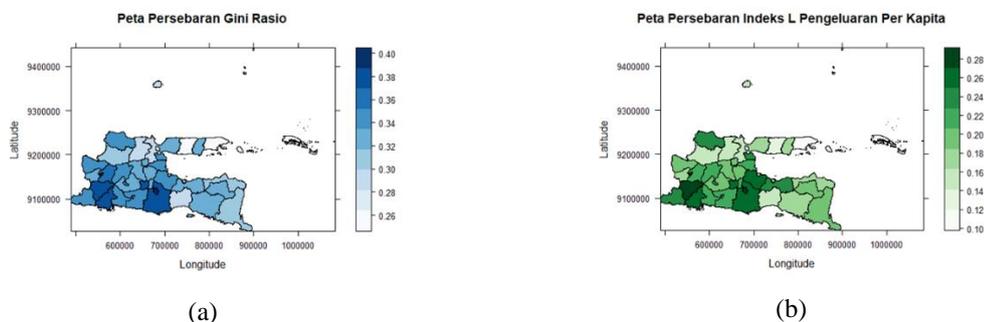
HASIL DAN PEMBAHASAN

Data pada penelitian ini merupakan data sekunder yang diambil melalui situs web resmi Badan Pusat Statistik (BPS) yaitu <https://www.bps.go.id/>. Penelitian ini menggunakan data pada tingkat kabupaten/kota di Provinsi Jawa Timur tahun 2020. Variabel respon *Y* pada penelitian ini adalah tingkat proporsi kasus AIDS dengan kategori 0 adalah kabupaten/kota dengan tingkat proporsi kasus AIDS yang rendah dan kategori 1 adalah kabupaten/kota dengan tingkat proporsi kasus AIDS tinggi. Variabel prediktor yang digunakan dalam penelitian ini antara lain gini rasio (GR), indeks L pengeluaran per kapita (IL.PPK), rasio jenis kelamin (SR), rasio ketergantungan (RK), indeks pembangunan *gender* (IPG), dan jumlah pos pelayanan KB desa (PPKB). Sebaran tingkat proporsi kasus AIDS di Provinsi Jawa Timur pada tahun 2020 dapat dilihat pada GAMBAR 1.



GAMBAR 2. Peta Persebaran Tingkat Proporsi Kasus AIDS Provinsi Jawa Timur tahun 2020

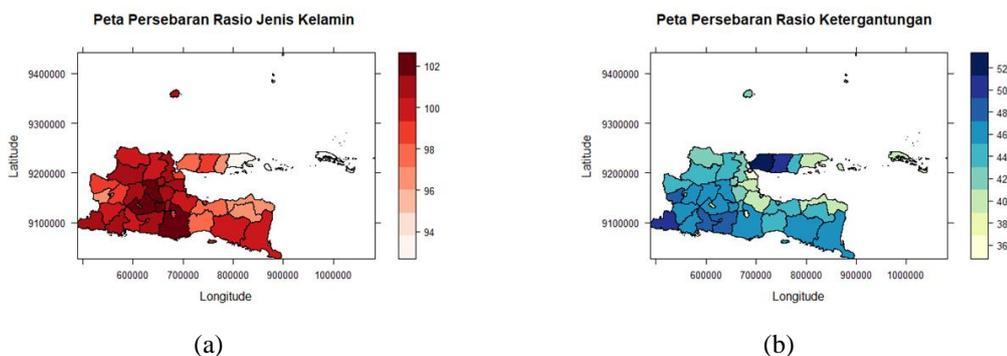
Berdasarkan pada GAMBAR 1, kabupaten/kota dengan tingkat proporsi kasus AIDS yang tinggi didominasi dengan kabupaten/kota yang memiliki jumlah penduduk yang tinggi seperti Kota Malang, Kota Madiun, dan Kota Surabaya, dimana Kota Surabaya merupakan kota dengan jumlah kasus AIDS paling banyak dan memiliki penduduk paling banyak di Provinsi Jawa Timur. Hal tersebut dapat menjelaskan bahwa Kota Surabaya berada di tingkat proporsi yang tinggi untuk kasus AIDS di Provinsi Jawa Timur. Kabupaten Pamekasan merupakan kabupaten dengan jumlah kasus AIDS paling sedikit pada tahun 2020 yaitu sebanyak 17 jiwa, tetapi memiliki jumlah penduduk yang cukup banyak yaitu sebesar 850.057 jiwa. Sehingga, Kabupaten Pamekasan menjadi kabupaten/kota dengan tingkat proporsi kasus AIDS yang rendah.



GAMBAR 3. Peta Persebaran (a) Gini Rasio; (b) Indeks L Pengeluaran Per Kapita

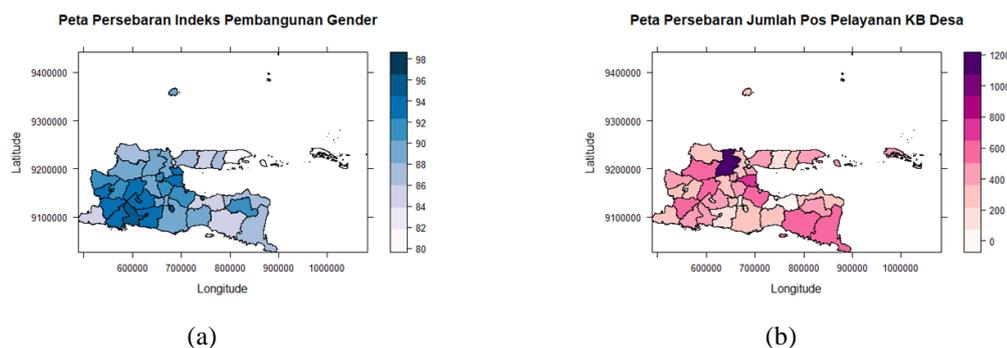
Berdasarkan GAMBAR 3(a), Kota Malang memiliki warna paling gelap yang menunjukkan bahwa ketimpangan pendapatan antar penduduk yang paling besar terjadi di Kota Malang. Sementara itu,

Kabupaten Sumenep memiliki warna yang paling terang di antara kabupaten/kota lainnya. Hal ini menandakan Kabupaten Sumenep memiliki ketimpangan pendapatan antar penduduk yang paling kecil. GAMBAR 3(b) menunjukkan semakin pekat warna di lokasi tersebut, maka semakin tinggi indeks L pengeluaran per kapitanya. Indeks L pengeluaran per kapita ini menunjukkan ketimpangan ekonomi dalam suatu kabupaten/kota. Ketimpangan ekonomi di Kabupaten Ponorogo dan Kabupaten Malang paling besar dibandingkan kabupaten/kota lainnya, sedangkan Kabupaten Sumenep memiliki ketimpangan ekonomi yang paling kecil.



GAMBAR 4. Peta Persebaran (a) Rasio Jenis Kelamin; (b) Rasio Ketergantungan

Pada GAMBAR 3(c) daerah dengan warna yang lebih gelap adalah kabupaten/kota yang memiliki rasio jenis kelamin lebih besar dari 100, yang menunjukkan bahwa penduduk laki-laki lebih banyak daripada perempuan. Kabupaten/kota tersebut antara lain Kabupaten Pacitan, Kabupaten Trenggalek, Kabupaten Tulungagung, Kabupaten Blitar, Kabupaten Kediri, Kabupaten Malang, Kabupaten Banyuwangi, Kabupaten Pasuruan, Kabupaten Sidoarjo, dsb. Kemudian, untuk GAMBAR 3(d) semakin gelap warna yang ditunjukkan pada lokasi pengamatan, semakin tinggi rasio ketergantungannya, yaitu semakin banyak beban tanggungan yang ditanggung oleh penduduk yang masih produktif. Kabupaten/kota Provinsi Jawa Timur yang berada di Pulau Jawa didominasi dengan warna-warna yang tidak gelap dengan interpretasi bahwa setiap 100 orang yang produktif memiliki tanggungan sekitar 36 - 50 orang.



GAMBAR 5. Peta Persebaran (a) Indeks Pembangunan Gender; (b) Jumlah Pos Pelayanan KB Desa

Semakin dekat nilai IPG dengan angka 100, maka semakin setara pembangunan antara wanita dan laki-laki. Berdasarkan GAMBAR 5(a), Kota Blitar merupakan lokasi dengan warna yang paling gelap dengan IPG sebesar 97,46. Sehingga, hal tersebut menandakan bahwa Kota Blitar adalah lokasi pengamatan dengan pembangunan yang setara antara penduduk wanita dan laki-laki. Semnetara, untuk GAMBAR 5(b) menunjukkan jumlah pos pelayanan KB, dimana merupakan jumlah kumulatif pos pembinaan dan pelayanan KB dari setiap desa yang berada di kabupaten/kota tersebut. Berdasarkan GAMBAR 5(b), Kabupaten Lamongan memiliki pos pelayanan KB paling banyak dengan jumlah 1.137 pos.

Sebelum memodelkan data perlu dilakukan pemeriksaan apakah terdapat multikolinieritas pada data. Pengujian multikolinieritas dapat menggunakan nilai *Variance Inflation Factor* (VIF). Nilai VIF dari variabel-variabel independen pada penelitian ini disajikan pada TABEL 2.

TABEL 2. Nilai VIF

| Variabel | Nilai VIF |
|--|-----------|
| Gini Rasio (GR) | 8,086908 |
| Indeks L Pengeluaran Per Kapita (IL.PKP) | 7,831353 |
| Rasio Jenis Kelamin (SR) | 1,762344 |
| Rasio Ketergantungan (RK) | 1,284920 |
| Indeks Pembangunan Gender (IPG) | 1,449023 |
| Jumlah Pos Pelayanan KB Desa (PPKB) | 1,537285 |

Berdasarkan TABEL 2, dapat dilihat bahwa nilai VIF dari seluruh variabel independent bernilai kurang dari 10. Sehingga, dapat disimpulkan bahwa tidak terdapat multikolinieritas antara variabel-variabel independen.

Setelah diperoleh bahwa tidak terdapat multikolinieritas pada data, maka akan diuji apakah terdapat heterogenitas spasial pada data. Pengujian heterogenitas pada data menggunakan uji *Breusch-Pagan* dengan hipotesis sebagai berikut:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{38}^2 = \sigma^2 \text{ (Tidak terdapat heterogenitas spasial)}$$

$$H_1: \text{Minimal terdapat satu } \sigma_i^2 \neq \sigma^2; i = 1, 2, \dots, 38 \text{ (Terdapat heterogenitas spasial)}$$

TABEL 3. Nilai Statistik Uji Heterogenitas Spasial

| BP | $\chi_{0,1;6}^2$ | P-value |
|--------|------------------|---------|
| 11,664 | 10,644 | 0,06991 |

Berdasarkan TABEL 3, dengan taraf signifikansi $\alpha = 0,1$ diperoleh nilai statistik uji $BP = 11,664 > \chi_{0,1;6}^2 = 10,644$ atau $p - value = 0,06991 < \alpha = 0,1$, maka H_0 ditolak. Sehingga, dapat disimpulkan bahwa terdapat heterogenitas spasial atau diduga terdapat pengaruh spasial pada data. Oleh karena itu, pemodelan dapat dilanjutkan menggunakan model *Geographically Weighted Logistic Regression* (GWLR).

Selanjutnya dilakukan pemodelan *Geographically Weighted Logistic Regression* (GWLR). Langkah awal dari melakukan penaksiran parameter model GWLR adalah menghitung jarak antar lokasi pengamatan berdasarkan *longitude* dan *latitude* menggunakan rumus jarak *Euclidean*. Setelah menghitung jarak antar lokasi pengamatan, maka akan ditentukan fungsi pembobot dan *bandwidth* optimum. Fungsi pembobot pada penelitian ini menggunakan *Fixed Gaussian* dengan *bandwidth* optimum ditentukan menggunakan *Akaike's Information Criterion Corrected* (AICc). Kemudian, dilakukan penaksiran parameter untuk model GWLR menggunakan *Maximum Likelihood Estimation* (MLE) dan diselesaikan dengan menggunakan metode iterasi numerik *Newton-Raphson*.

Pengujian kesesuaian model dilakukan untuk melihat apakah terdapat perbedaan antara model regresi logistik biner global dengan model *Geographically Weighted Logistic Regression* (GWLR) serta melihat model mana yang lebih representatif atau lebih sesuai untuk data penelitian. Kesesuaian model dilihat berdasarkan ketidaksamaan model dan berdasarkan nilai AIC. Hipotesis untuk pengujian kesamaan model sebagai berikut:

$$H_0: \beta_k(u_i, v_i) = \beta_k; i = 1, 2, \dots, 38; k = 1, 2, \dots, 6 \text{ (Tidak terdapat perbedaan antara model global dan model GWLR)}$$

$$H_1: \text{Minimal terdapat satu } \beta_k(u_i, v_i) \neq \beta_k \text{ (Terdapat perbedaan antara model global dan model GWLR)}$$

TABEL 4. Nilai Statistik Uji Kesesuaian Model

| $D(\beta)$ | df_1 | $D(\beta^*)$ | df_2 | F_{hit} | $F_{0,1;6;228}$ |
|------------|--------|--------------|--------|-----------|-----------------|
| 31,441186 | 6 | 27,530708 | 228 | 43,397542 | 0,37 |

Berdasarkan TABEL 4, dengan menggunakan tingkat signifikansi $\alpha = 0,1$ diperoleh $F_{hit} = 43,397542 > F_{0,1;6;228} = 0,37$. Keputusan dari uji kesesuaian model tersebut adalah H_0 ditolak. Sehingga, dapat disimpulkan bahwa model regresi logistik biner global dan model *Geographically Weighted Logistic Regression* (GWLR) tidak identik atau terdapat perbedaan di antara kedua model tersebut.

Berdasarkan uji kesamaan model, diperoleh bahwa model Regresi Logistik Biner Global dan model *Geographically Weighted Logistic Regression* (GWLR) tidak identik. Kedua model tersebut akan dibandingkan dengan menggunakan nilai *Akaike Information Criterion* (AIC) untuk melihat model manakah yang lebih sesuai dalam menjelaskan data penelitian.

TABEL 5. Nilai AIC

| Model | Nilai AIC |
|-------------------------------|-----------|
| Regresi Logistik Biner Global | 45,441186 |
| GWLR <i>Fixed Gaussian</i> | 44,177951 |

Berdasarkan TABEL, 5 diperoleh bahwa nilai AIC model GWLR lebih kecil dibandingkan dengan model Regresi Logistik Biner Model. Hal tersebut menunjukkan bahwa model GWLR dengan fungsi pembobot kernel *Fixed Gaussian* lebih sesuai dalam menjelaskan faktor-faktor yang memengaruhi tingkat proporsi kasus AIDS Provinsi Jawa Timur tahun 2020.

Setelah menguji kesesuaian model, akan dilakukan uji signifikansi parameter. Pengujian parameter secara serentak dilakukan untuk mengetahui apakah variabel-variabel independen terdapat pengaruh yang signifikan terhadap variabel respon secara serentak. Hipotesis pengujian parameter secara serentak untuk model GWLR sebagai berikut:

$$H_0: \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_k(u_i, v_i) = 0; k = 1,2, \dots, 6; i = 1,2, \dots, 38$$

$$H_1: \text{Minimal terdapat satu } \beta_k(u_i, v_i) \neq 0; k = 1,2, \dots, 6$$

TABEL 6. Nilai Statistik Uji Signifikansi Parameter Secara Serentak

| G_2 | $\chi^2_{\alpha, v}$ |
|-----------|----------------------|
| 27,530708 | 13,31156614 |

Berdasarkan TABEL 6, dengan menggunakan tingkat signifikansi $\alpha = 0,1$, diperoleh nilai statistik uji $G_2 = 27,530708 > \chi^2_{0,1;8,3} = 13,31156614$. Hal tersebut menunjukkan bahwa hasil keputusan pengujian parameter model GWLR secara serentak adalah H_0 ditolak. Sehingga, dapat dikatakan bahwa setidaknya terdapat satu variabel independen yang signifikan terhadap variabel dependennya di setiap lokasi pengamatan.

Pengujian parameter secara parsial bertujuan untuk mengetahui apakah variabel-variabel independen terdapat pengaruh yang signifikan terhadap variabel respon secara individual. Hipotesis pengujian parameter secara parsial untuk model GWLR sebagai berikut:

$$H_0: \beta_k(u_i, v_i) = 0; k = 1,2, \dots, 6; i = 1,2, \dots, 38$$

$$H_1: \text{Minimal terdapat satu } \beta_k(u_i, v_i) \neq 0; k = 1,2, \dots, 6$$

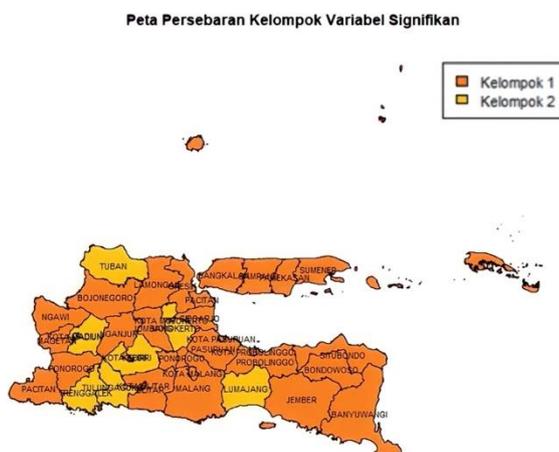
Hasil pengujian parameter secara parsial digambarkan dalam kelompok variabel-variabel prediktor yang secara signifikan memengaruhi variabel respon. Berikut kelompok model GWLR berdasarkan variabel-variabel prediktor yang signifikan:

TABEL 7. Kelompok Model GWLR Berdasarkan Variabel-Variabel Prediktor yang Signifikan

| Kelompok | Variabel yang Signifikan | Lokasi Pengamatan |
|----------|--|--|
| 1 | Rasio Jenis Kelamin dan Rasio Ketergantungan | Kab. Pacitan, Kab. Ponorogo, Kab. Blitar, Kab. Malang, Kab. Jember, Kab. Banyuwangi, Kab. Bondowoso, Kab. Situbondo, Kab. Probolinggo, Kab. Pasuruan, Kab. Sidoarjo, Kab. Jombang, Kab. Nganjuk, Kab. Magetan, Kab. Ngawi, Kab. Bojonegoro, Kab. Gresik, Kab. Bangkalan, Kab. Sampang, Kab. Pamekasan, Kab. Sumenep, Kota Kediri, Kota Blitar, Kota Malang, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu |
| 2 | Rasio Ketergantungan | Kab. Trenggalek, Kab. Tulungagung, Kab. Kediri, Kab. Lumajang, Kab. Mojokerto, Kab. Madiun, Kab. Tuban, Kota Madiun |

Kelompok kabupaten/kota berdasarkan variabel signifikannya secara visualisasi dapat dilihat pada GAMBAR 2. Berdasarkan TABEL 7 dan GAMBAR 2, 76,5% tingkat proporsi kasus AIDS kabupaten/kota di Provinsi Jawa Timur yang berada di daerah Pulau Jawa dipengaruhi oleh rasio jenis kelamin dan rasio ketergantungan, 23,5% sisanya hanya dipengaruhi oleh rasio ketergantungan saja. Sedangkan untuk kabupaten/kota Provinsi Jawa Timur yang berada di daerah Pulau Madura 100% tingkat proporsi kasus AIDSnya dipengaruhi oleh rasio jenis kelamin dan rasio ketergantungan.

Model yang paling sesuai dalam menjelaskan faktor-faktor pada data penelitian adalah model GWLR. Sehingga, nilai Z hitung yang digunakan adalah nilai Z hitung parameter model GWLR, dengan nilai Z hitung adalah estimasi parameter dibagi dengan standar error. Pembagian kluster diperoleh berdasarkan *centroid*, nilai titik-titik yang lebih dekat dengan masing-masing *centroid* akan berada di dalam kluster yang sama. *Centroid* masing-masing kluster disajikan pada TABEL 8.



GAMBAR 6. Peta Persebaran Kelompok Berdasarkan Variabel-Variabel yang Signifikan

TABEL 8. Centroid Analisis Kluster K-Means Berdasarkan Nilai Z Hitung

| Kluster | Z Hitung Estimasi Parameter | | | | | |
|---------|-----------------------------|-----------|-----------|-----------|-----------|-----------|
| | β_1 | β_2 | β_3 | β_4 | β_5 | β_6 |
| 1 | 0,1264 | 0,4979 | 2,0453 | 2,2053 | 0,3219 | 1,3988 |
| 2 | 0,4539 | 1,1575 | 1,4087 | 1,8255 | 0,6347 | 1,1126 |

Hasil pengelompokan kabupaten/kota berdasarkan nilai Z hitung menggunakan kluster *k-means* disajikan pada TABEL 9.

TABEL 9. Hasil Pengelompokan Klaster *K-Means*

| Klaster | Lokasi Pengamatan |
|----------------|---|
| 1 | Kab. Bangkalan, Kab. Banyuwangi, Kab. Blitar, Kab. Bojonegoro, Kab. Bondowoso, Kab. Gresik, Kab. Jember, Kab. Jombang, Kota Batu, Kota Blitar, Kota Kediri, Kota Malang, Kota Mojokerto, Kota Pasuruan, Kota Probolinggo, Kota Surabaya, Kab. Malang, Kab. Nganjuk, Kab. Ngawi, Kab. Pacitan, Kab. Pamekasan, Kab. Pasuruan, Kab. Ponorogo, Kab. Probolinggo, Kab. Sampang, Kab. Sidoarjo, Kab. Situbondo, Kab. Sumenep |
| 2 | Kab. Kediri, Kota Madiun, Kab. Lamongan, Kab. Lumajang, Kab. Madiun, Kab. Magetan, Kab. Mojokerto, Kab. Trenggalek, Kab. Tuban, Kab. Tulungagung |

Berdasarkan TABEL 9, 74% tingkat proporsi kasus AIDS di kabupaten/kota Provinsi Jawa Timur cenderung dipengaruhi oleh rasio jenis kelamin dan rasio ketergantungan, sementara 26% sisanya hanya dipengaruhi oleh rasio ketergantungan saja. Tingkat proporsi kasus AIDS kabupaten/kota yang hanya dipengaruhi oleh rasio ketergantungan hanya terdapat pada daerah yang berada di Pulau Jawa, daerah yang berada di Pulau Madura 100% dipengaruhi oleh rasio jenis kelamin dan rasio ketergantungan. Dalam hal ini, adanya pengaruh signifikan variabel rasio jenis kelamin menandakan bahwa semakin tinggi jumlah penduduk laki-laki di kabupaten/kota tersebut, akan meningkatkan resiko kabupaten/kota tersebut memiliki tingkat proporsi kasus AIDS yang tinggi. Sedangkan untuk kabupaten/kota yang dipengaruhi secara signifikan oleh variabel rasio ketergantungan menandakan semakin tinggi jumlah penduduk yang belum produktif dan tidak produktif lagi, maka akan meningkatkan resiko kabupaten/kota tersebut memiliki tingkat proporsi kasus AIDS yang tinggi.

KESIMPULAN DAN SARAN

Berdasarkan model *Geographically Weighted Logistic Regression* (GWLRL), diperoleh variabel-variabel yang signifikan untuk setiap daerahnya. Terdapat 2 kelompok daerah berdasarkan variabel-variabel signifikannya, yaitu kelompok 1 dengan variabel signifikannya adalah rasio jenis kelamin dan rasio ketergantungan, antara lain Kab. Pacitan, Kab. Ponorogo, Kab. Blitar, Kab. Malang, Kab. Jember, Kab. Banyuwangi, Kab. Bondowoso, Kab. Situbondo, Kab. Probolinggo, Kab. Pasuruan, Kab. Sidoarjo, Kab. Jombang, Kab. Nganjuk, Kab. Magetan, Kab. Ngawi, Kab. Bojonegoro, Kab. Gresik, Kab. Bangkalan, Kab. Sampang, Kab. Pamekasan, Kab. Sumenep, Kota Kediri, Kota Blitar, Kota Malang, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota Batu. Sementara, untuk kelompok 2 dengan variabel signifikannya hanya rasio ketergantungan, antara lain Kab. Trenggalek, Kab. Tulungagung, Kab. Kediri, Kab. Lumajang, Kab. Mojokerto, Kab. Madiun, Kab. Tuban, dan Kota Madiun.

Berdasarkan hasil pengelompokan menggunakan analisis kluster *k-means*, diperoleh 2 klaster berdasarkan nilai *Z* hitung. Berdasarkan nilai *centroid*, kabupaten/kota anggota klaster 1 memiliki kecenderungan merupakan lokasi pengamatan dengan variabel signifikannya yaitu rasio jenis kelamin (SR) dan rasio ketergantungan (RK), sedangkan kabupaten/kota anggota klaster 2 memiliki kecenderungan merupakan lokasi pengamatan dengan variabel signifikannya yaitu rasio ketergantungan (RK). Adanya pengaruh signifikan variabel rasio jenis kelamin menandakan bahwa semakin tinggi jumlah penduduk laki-laki di kabupaten/kota tersebut, akan meningkatkan resiko kabupaten/kota tersebut memiliki tingkat proporsi kasus AIDS yang tinggi. Sedangkan untuk kabupaten/kota yang dipengaruhi secara signifikan oleh variabel rasio ketergantungan menandakan semakin tinggi jumlah penduduk yang belum produktif dan tidak produktif lagi, maka akan meningkatkan resiko kabupaten/kota tersebut memiliki tingkat proporsi kasus AIDS yang tinggi.

Adapun saran yang direkomendasikan yaitu penelitian selanjutnya dapat menemukan faktor-faktor lain yang memengaruhi tingginya kasus AIDS. Selain itu, melakukan penelitian di provinsi-provinsi lainnya agar pemerintah di setiap provinsi dapat mengurangi jumlah kasus AIDS yang ada di Indonesia.

UCAPAN TERIMA KASIH

Penelitian ini didukung oleh Direktorat Riset dan Pengembangan Universitas Indonesia, melalui Program Hibah Publikasi Terindeks Internasional (PUTI) Q2 Tahun Anggaran 2022-2023 Nomor: NKB-668/UN2.RST/HKP.05.00/2022 sebagai luaran tambahan.

REFERENSI

- Abonyi, J. dan Feil, B. (2007). *Cluster Analysis for Data Mining and System Identification*. Basel: Birkhasuer Verlag AG.
- Agresti, A. (2013). *Categorical Data Analysis*. New Jersey: John Wiley & Sons.
- Badan Pusat Statistik. (2022). Gini Rasio. Tersedia di: <https://sirusa.bps.go.id/sirusa/index.php/indikator/999> (Diakses: 23 Maret 2022).
- Badan Pusat Statistik. (2022). Indeks Pembangunan Gender. Tersedia di: <https://sirusa.bps.go.id/sirusa/index.php/indikator/14> (Diakses: 23 Maret 2022).
- Badan Pusat Statistik. (2022). Pengeluaran Per Kapita. Tersedia di: <https://sirusa.bps.go.id/sirusa/index.php/indikator/197> (Diakses: 23 Maret 2022).
- Badan Pusat Statistik. (2022). Rasio Jenis Kelamin. Tersedia di: <https://sirusa.bps.go.id/sirusa/index.php/indikator/87> (Diakses: 23 Maret 2022).
- Badan Pusat Statistik. (2022). Rasio Ketergantungan. Tersedia di: <https://sirusa.bps.go.id/sirusa/index.php/indikator/95> (Diakses: 23 Maret 2022).
- Caraka, R.E. dan Yasin, H. (2017). *Geographically Weighted Regression: Sebuah Kajian Geografis*. Yogyakarta: Mobius.
- Fotheringham, A. S., Brunson, C., & Charlton, M. (2002). *Quantitative Geography: Perspective on Spatial Data Analysis*. London: SAGE Publication.
- Hidayat, S. dan Woyanti, N. (2021). Pengaruh PDRB Per Kapita, Belanja Daerah, Rasio Ketergantungan, Kemiskina, dan Teknologi Terhadap IPM Indonesia. *Jurnal Ekonomi, Bisnis, dan Akuntansi*, Vol. 3, No. 4, hh. 132-136.
- Hosmer Jr., D. W., Lemeshow, S., dan Sturdivant, R. X. (2013). *Applied Logistic Regression*, Third Edition. New Jersey: John Wiley & Sons.
- Kementerian Kesehatan Republik Indonesia. (2021). *Infodatin AIDS*. Jakarta: Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia.
- Mendenhall, W. dan Sincich, T. (2012) *A Second Course in Statistics Regression Analysis*. Boston: Pearson Education.
- National Institute of Health. (2021). *The Stage of Infection*. Tersedia di: <https://hivinfo.nih.gov/understanding-hiv/fact-sheets/stages-hiv-infection> (Diakses: 20 Maret 2022).
- Ratnasari, N. T. dan Purhadi. (2013). Pemodelan Faktor yang Mempengaruhi Jumlah Hiv dan Aids Provinsi Jawa Timur Menggunakan Regresi Poisson Bivariat. *Jurnal Sains dan Seni Pomits*, Vol. 2, No. 2, hh. 213-218.
- Wu, Junjie. (2012). *Advances in K-means Clustering: A Data Mining Thinking*. Berlin: Springer.