

Pemodelan Faktor-Faktor yang Berpengaruh Terhadap Angka Buta Huruf Di Provinsi Jawa Barat dengan *Geographically Weighted Logistic Regression*

¹Amanda Tiara Pravitasary, ²Nusar Hajarisman, ³Siti Sunendiari

*Program Studi Statistika, Universitas Islam Bandung,
Jl. Tamansari No. 1 Bandung 40116*

e-mail:¹ amandatiaraa@gmail.com

Abstrak. Analisis regresi logistik merupakan analisis yang digunakan untuk mengetahui bagaimana hubungan antara peubah respon yang bersifat diskrit dengan satu atau lebih peubah prediktor. Analisis ini akan menghasilkan model yang kurang tepat apabila diterapkan pada data yang dipengaruhi lokasi secara geografis atau biasa disebut data spasial. Hal ini karena model yang diperoleh oleh regresi logistik tidak mampu menangkap efek lokal dari karakteristik yang ditimbulkan oleh masing-masing lokasi tersebut. Metode yang dapat digunakan pada data spasial adalah *Geographically Weighted Logistic Regression* (GWLR), yang merupakan bentuk lokal dari regresi logistik dimana karakteristik dari masing-masing lokasi dipertimbangkan. Dalam membentuk model GWLR diperlukan fungsi pembobot yang dapat dihitung dengan menyertakan nilai jarak antar lokasi dan nilai *bandwidth*. Perbandingan model regresi logistik dan model GWLR dapat dilakukan dengan melihat nilai AIC terkecil diantara keduanya, hal ini bertujuan untuk mendapatkan model terbaik.

Kata Kunci: *Regresi Logistik, GWLR, Fungsi Pembobot, Bandwidth, AIC*

A. Pendahuluan

Regresi merupakan salah satu bagian analisis data statistika yang digunakan untuk mendapatkan model hubungan antara variabel respon (Y) dengan satu atau lebih variabel prediktor (X). Model dari regresi linear biasa memiliki variabel respon bersifat kontinu dan diasumsikan mengikuti distribusi normal. Akan tetapi terkadang juga ditemukan variabel respon yang bersifat diskrit maka model regresi yang dapat digunakan adalah model regresi logistik.

Sebagaimana model regresi linier biasa, model regresi logistik pada umumnya digunakan ketika unit pengamatannya bukan merupakan wilayah atau lokasi, karena model tersebut tidak memperhitungkan efek wilayah atau lokasi yang selanjutnya disebut efek spasial. Analisis regresi global (seperti regresi biasa maupun regresi logistik) bukan merupakan analisis data yang tepat jika digunakan pada data spasial. Masalah utama jika analisis ini diterapkan pada data yang mengandung informasi lokasi geografis atau spasial, maka akan terjadi heterogenitas spasial. Terjadinya heterogenitas spasial pada parameter regresi menyebabkan analisis regresi global tidak mampu menangkap efek lokal dari karakteristik yang ditimbulkan oleh masing-masing lokasi. Oleh karena itu dikembangkan sebuah metode yang dapat digunakan pada data spasial yaitu *Geographically Weighted Regression* (GWR). Pada beberapa kasus, metode GWR tidak dapat memberikan hasil yang lebih baik dibandingkan dengan analisis regresi global, terutama apabila tidak terjadi masalah heterogenitas. Melalui GWR akan dibentuk model pada kumpulan data yang bergantung pada lokasi atau letak secara geografis dimana data tersebut diamati (Brunsdon, *et. al.*, 2002). Metode GWR yang merupakan pengembangan dari regresi global untuk variabel respon yang bersifat kontinu dikembangkan untuk memprediksi atau menduga model dari kumpulan data yang memiliki peubah respon biner melalui model logistik. Metode ini disebut dengan *Geographically Weighted Logistic Regression* (GWLR) (Atkinson *et. al.*, 2003).

Metode GWR telah banyak dilakukan di Indonesia, salah satunya metode ini dapat diaplikasikan untuk menyelidiki variabel-variabel yang berpengaruh terhadap Angka Buta Huruf (ABH). Penelitian ABH antara lain pernah dilakukan oleh Lailiyah dan Purhadi (2012) dengan *Geographically Weighted Ordinal Logistic Regression* (GWOLR) untuk memodelkan faktor-faktor yang mempengaruhi tingkat buta huruf Kabupaten/Kota di Jawa Timur. Metode-metode yang lain yang diaplikasikan pada Angka Buta Huruf antara lain dilakukan oleh Eka dan Sutikno (2011) dengan judul pemodelan dan pemetaan angka buta huruf Provinsi Jawa Timur dengan pendekatan regresi spasial, dan Giovani *et. al.*, (2013) yaitu pemodelan faktor-faktor yang mempengaruhi Angka Buta Huruf (ABH) Kabupaten/Kota di Jawa Timur dengan Regresi Spline Semiparametrik.

Di Provinsi Jawa Barat, buta huruf merupakan salah satu permasalahan yang paling mendasar di masyarakat yang masih belum teratasi. Tercatat bahwa Angka Melek Huruf (AMH) di tahun 2012 adalah 96.18% (BPS, 2012). Untuk itu berdasarkan permasalahan yang telah dibahas maka dalam skripsi ini akan dibahas mengenai model regresi logistik dan model regresi yang mampu menangkap efek lokal, dimana variabel respon merupakan data biner melalui metode GWLR. Model GWLR ini akan diterapkan pada faktor-faktor yang mempengaruhi tingkat ABH Kabupaten/Kota di Jawa Barat.

B. Tinjauan Pustaka

Model Regresi Logistik

Regresi logistik merupakan salah satu metode yang dapat digunakan untuk mencari hubungan variabel respon yang bersifat dikotomis atau biner dengan satu atau lebih variabel prediktor yang bersifat kontinu (Hosmer dan Lemeshow, 2000). Variabel respon dari regresi logistik biner dapat dinyatakan dengan dua nilai yaitu 0 dan 1 dimana pengkategorian ini mengikuti distribusi Bernoulli dengan distribusi peluangnya adalah :

$$P(Y = y_i) = \pi^{y_i} (1 - \pi)^{1 - y_i} \quad \text{untuk } y_i = 0, 1 \quad \dots(2.1)$$

Secara umum model probabilitas regresi logistik dimana $\pi(x)$ bergantung pada variabel prediktor $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i1}, \dots, x_{ip})^T$ dapat menggunakan persamaan sebagai berikut (Hosmer dan Lemeshow, 2000):

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})} \quad \dots(2.2)$$

Bentuk $\pi(x)$ tersebut kemudian dilakukan transformasi logit bentuk persamaannya adalah sebagai berikut:

$$\text{logit}(\pi(x)) = \log \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} = \mathbf{x}_i^T \boldsymbol{\beta} \quad \dots(2.3)$$

Penaksir parameter-parameter dalam model logistik dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Pada dasarnya metode maksimum likelihood memberikan nilai penaksir $\boldsymbol{\beta}$ dengan memaksimalkan fungsi *likelihood* (Hosmer dan Lemeshow, 2000). Oleh karena model regresi logistik merupakan model non linear maka, penaksir kemungkinan maksimum harus diselesaikan secara numerik melalui metode Newton Rapsion.

Model GWLR

Geographically Weighted Logistic Regression (GWLR) merupakan salah satu metode regresi yang dapat mempertimbangkan faktor spasial sehingga akan dihasilkan

nilai parameter bagi masing-masing titik atau lokasi dimana data tersebut diamati. Metode ini dikembangkan dari metode GWR yang digunakan untuk memprediksi atau menduga model dari kumpulan data yang memiliki peubah respon biner melalui model logistik (Atkinson *et. al*, 2003).

Pada teknik GWLR lokasi geografis dimasukkan ke dalam model melalui fungsi pembobot. Pembobot (w_{ij}) diberikan pada masing-masing observasi. Sehingga model yang terbentuk adalah:

$$\pi(x_i) = \frac{\exp(\beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \dots + \beta_p(u_i, v_i)x_{ip})}{1 + \exp(\beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \dots + \beta_p(u_i, v_i)x_{ip})} \quad \dots(2.4)$$

Bentuk logit untuk model GWLR yang dinyatakan dengan $g(x_i)$ adalah sebagai berikut (Brundson *et. al*, 2002):

$$\text{logit } \pi(x_i) = \log \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \dots + \beta_p(u_i, v_i)x_{ip} \quad \dots(2.5)$$

dimana :

- ~ $\beta_0(u_i, v_i)$ = konstanta (intersep) pada masing-masing lokasi.
- ~ $\beta_1(u_i, v_i), \beta_2(u_i, v_i), \dots, \beta_p(u_i, v_i)$ = koefisien regresi ke-1, 2, ..., p pada masing-masing lokasi.
- ~ $X_{i1}, X_{i2}, \dots, X_{ip}$ = variabel prediktor ke-1, 2, ..., p pada lokasi ke-i
- ~ (u_i, v_i) = koordinat *longitude-latitude* dari titik ke-i pada suatu lokasi geografis.

Penaksiran parameter pada model GWLR melibatkan suatu fungsi pembobot. Peran fungsi pembobot sangat bergantung pada jarak antar titik lokasi pengamatan dan nilai *bandwidth* yang optimum. Perhitungan *bandwidth* yang optimum salah satunya dapat digunakan melalui *Akaike Information Criterion* (AIC). Adapun persamaannya adalah sebagai berikut :

$$AIC = 2n \log_e(\hat{\sigma}) + n \log_e(2\pi) + n + tr(S) \quad \dots(2.6)$$

Selain untuk menentukan *bandwidth* optimum, AIC dapat juga digunakan untuk menentukan model terbaik.

Ada beberapa macam fungsi pembobot yang dapat digunakan salah satunya adalah fungsi pembobot *Fixed Bisquare Kernel*. Fungsi ini akan memiliki nilai *bandwidth* yang sama pada setiap lokasi. Menurut Brundson *et.al*, (2002), fungsi pembobotnya dapat dituliskan sebagai berikut:

$$w_{ij}(u_i, v_i) = \begin{cases} [1 - (d_{ij} / h)^2]^2, & \text{jika } d_{ij} \leq h \\ 0, & \text{jika } d_{ij} > h \end{cases} \quad \dots(2.7)$$

keterangan : d_{ij} adalah jarak *euclidean* antara lokasi i terhadap lokasi j
 h adalah nilai *bandwith* optimum

Dalam Anggarini dan Purhadi (2012) penaksiran parameter dalam model GWLR adalah menggunakan MLE terboboti. Adapun fungsi *likelihood* yang terbentuk adalah

$$L(\beta(u_i, v_i)) = \left\{ \prod_{i=1}^n \left[1 + \exp \sum_{j=0}^p \beta_j(u_i, v_i)x_{ij} \right]^{-1} \right\} \exp \left[\sum_{j=0}^p \left(\sum_{i=1}^n y_i x_{ij} \right) \beta_j(u_i, v_i) \right] \quad \dots(2.8)$$

dan logaritma *likelihood* dapat dinyatakan sebagai:

$$\ln L(\beta(u_i, v_i)) = \sum_{j=0}^p \left(\sum_{i=1}^n y_i x_{ij} \right) \beta_j(u_i, v_i) - \sum_{i=1}^n \ln \left\{ 1 + \exp \left(\sum_{j=0}^p \beta_j(u_i, v_i)x_{ij} \right) \right\} \quad \dots(2.9)$$

Taksiran parameter β dapat diperoleh dari metode maksimum *likelihood* terboboti. Karena model dari GWLR merupakan model non linear maka selanjutnya digunakan metode Newton- Rhapson untuk menaksir parameter-parameternya.

Pengujian parameter model GWLR dilakukan secara serentak dan secara parsial, Pengujian secara serentak hipotesis-nya adalah sebagai berikut:

$$H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \dots = \beta_j(u_i, v_i) = 0; j = 1, 2, \dots, p; i = 1, 2, \dots, n$$

$$H_1 : \text{paling tidak terdapat satu } \beta_j(u_i, v_i) \neq 0$$

Dengan statistik ujinya (Hosmer dan Lemeshow, 2000) adalah

$$G^2(u_i, v_i) = -2 \sum_{i=1}^n \sum_{k=1}^K y_{ij} \ln \left(\frac{\sum_{j=1}^p y_{jk} w_j(u_i, v_i)}{\sum_{j=1}^p w_j(u_i, v_i)} \right) - 2 \sum_{i=1}^n \sum_{k=1}^K y_{ik} \ln(\pi(x_i)) \quad \dots(2.10)$$

Aturan keputusan :

H_0 ditolak jika $G^2 > \chi^2_{(v, \alpha)}$, yang artinya paling tidak terdapat satu $\beta_k(u_i, v_i) \neq 0$

Uji parameter secara parsial digunakan untuk mengetahui parameter yang berpengaruh secara signifikan terhadap model. Hipotesis dari uji parsial adalah sebagai berikut:

$$H_0 : \beta_k(u_i, v_i) = 0 \quad ; k=0, 1, 2, \dots, p$$

$$H_1 : \beta_k(u_i, v_i) \neq 0$$

Statistik uji yang digunakan (Hosmer dan Lemeshow, 2000) adalah:

$$W = \frac{\hat{\beta}_k(u_i, v_i)}{SE(\hat{\beta}_k(u_i, v_i))} \quad \dots(2.11)$$

aturan keputusan : H_0 ditolak jika $|W| > Z_{\alpha/2}$

C. Hasil dan Pembahasan

Bahan yang digunakan adalah data sekunder yang merupakan hasil pencatatan Badan Pusat Statistika (BPS) pada tahun 2012 yang diperoleh dari katalog Indeks Pembangunan Manusia, Indikator Kesejahteraan Rakyat, Jawa Barat Dalam Angka dan Data Informasi Kemiskinan Kabupaten/Kota.

Dalam upaya penurunan ABH, Kementerian Pendidikan dan Kebudayaan (Kemendikbud) menentukan target ABH tahun 2012 adalah sebesar 4.8 % penduduk (www.kemendikbud.go.id). Variabel respon pada penelitian ini adalah ABH yang berada di bawah 4.8 % dan ABH yang berada di atas 4.8 %. Variabel prediktor yang digunakan adalah Tingkat Pengangguran Terbuka (X_1), persentase penduduk miskin (X_2), Angka Partisipasi Sekolah Sekolah Dasar (X_3) dan persentase daerah berstatus kota (X_4). Selain itu juga, pada penelitian ini diperlukan data mengenai koordinat spasial (*longitude* dan *latitude*) tiap kabupaten/kota di Provinsi Jawa Barat.

Model Regresi Logistik

Model peluang Provinsi Jawa Barat dalam menekan ABH untuk berada di bawah target 4.8% dengan melibatkan seluruh variabel prediktor disajikan dalam Persamaan (3.1):

$$\pi(x) = \frac{\exp(169.2690 - 0.3895 x_1 - 0.1322 x_2 - 1.6821x_3 + 0.0502x_4)}{1 + \exp(169.2690 - 0.3895x_1 - 0.1322 x_2 - 1.6821x_3 + 0.0502 x_4)} \quad \dots(3.1)$$

Pada Persamaan (3.1), Pengujian koefisien regresi secara serentak dilakukan dengan menggunakan uji rasio kemungkinan (G^2) diperoleh nilai $G^2_{hitung} = 18.717$ yang lebih besar dari $\chi^2_{(0.05,4)} = 11.1$. Hal ini menunjukkan bahwa paling tidak terdapat satu $\beta \neq 0$ atau minimal ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon. Selain itu dengan melakukan pengujian secara parsial dengan menggunakan tingkat signifikansi (α) sebesar 10%, terdapat dua variabel prediktor yang berpengaruh signifikan yaitu $APSD$ dengan nilai $Z_{wald} = -1.671$ dan Daerah Berstatus Kota dengan nilai $Z_{wald} = 1.842$. Hal tersebut ditunjukkan oleh nilai Z_{wald} yang lebih besar dari $Z_{\alpha/2} = 1.64$.

Variabel yang signifikan yaitu X_3 dan X_4 kemudian di regresikan kembali dengan variabel Y , sehingga model regresi logistik dengan X_3 dan X_4 adalah :

$$\pi(x) = \frac{\exp(128.4812 - 1.3126x_3 + 0.0388 x_4)}{1 + \exp(128.4812 - 1.3126x_3 + 0.0388 x_4)} \quad \dots(3.2)$$

Model GWLR

Pada pemodelan GWLR ini akan dihasilkan 26 buah model untuk setiap kabupaten/kota di Provinsi Jawa Barat. Langkah-langkah dalam pemodelan GWLR adalah menentukan jarak *euclidean*, *bandwidth* optimum, pembobot dan penaksiran parameter GWLR. Penentuan *bandwidth* optimum pada penelitian ini dengan bantuan software *GWR4* diperoleh nilai AIC yang minimum adalah 28.797 dengan nilai *bandwidth* optimumnya adalah 28.647.

Pengujian secara serentak pada model GWLR diperoleh nilai statistik uji G^2 adalah 18.677. Sehingga dengan tingkat signifikan (α) = 5% diperoleh nilai $\chi^2_{(0.05,4)} = 11.1$, maka dapat diputuskan bahwa paling tidak ada satu variabel prediktor yang berpengaruh signifikan terhadap variabel respon pada masing-masing kabupaten/kota di Jawa Barat. Hal ini dapat ditunjukkan dengan statistik uji G^2 yang lebih besar dari $\chi^2_{(0.05,4)}$. Pengujian koefisien regresi secara parsial pada masing-masing wilayah dengan menggunakan tingkat signifikan pada $\alpha = 10\%$ diperoleh nilai $z(\alpha/2) = 1.64$, hasilnya disimpulkan bahwa ada dua variabel yaitu variabel $APSD$ (X_3) dan variabel presentase daerah berstatus kota (X_4) yang berpengaruh signifikan untuk menekan ABH di masing-masing kota/kabupaten di Jawa Barat.

Variabel yang signifikan kemudian di regresikan kembali, sehingga model GWLR dengan melibatkan variabel yang signifikan yaitu X_3 dan X_4 hasilnya disajikan pada Tabel 3.1 :

Tabel 3.1

Hasil Penaksiran Parameter Model GWLR dengan X_3 dan X_4

Kota / Kabupaten	Konstanta	β_3	β_4	$\pi(x)$
Kab Bogor	128.83775	-1.31627	0.03891	0.9373
Kab Sukabumi	128.85919	-1.31646	0.03888	0.8796
Kab Cianjur	128.90276	-1.31692	0.03891	0.6010
Kab Bandung	128.52378	-1.31303	0.03881	0.8031
Kab Garut	128.34370	-1.31117	0.03874	0.8104
Kab Tasikmalaya	128.45365	-1.31229	0.03876	0.8017
Kab Ciamis	128.28942	-1.31060	0.03871	0.6464
Kab Kuningan	128.19289	-1.30964	0.03872	0.5591
Kab Cirebon	128.11326	-1.30884	0.03872	0.9102
Kab Majalengka	128.28802	-1.31062	0.03875	0.3296
Kab Sumedang	128.47073	-1.31249	0.03880	0.5128
Kab Indramayu	128.27955	-1.31056	0.03878	0.8626
Kab Subang	128.48456	-1.31266	0.03883	0.3214
Kab Purwakarta	128.54242	-1.31325	0.03884	0.9464
Kab Karawang	128.51144	-1.31295	0.03885	0.5185
Kab Bekasi	128.68133	-1.31469	0.03889	0.6840
Kab Bandung Barat	128.52378	-1.31303	0.03881	0.9815
Kota Bogor	128.83131	-1.31621	0.03892	0.9481
Kota Sukabumi	128.88551	-1.31675	0.03891	0.9653
Kota Bandung	128.79957	-1.31586	0.03888	0.9301
Kota Cirebon	128.11326	-1.30884	0.03872	0.9607
Kota Bekasi	128.69447	-1.31482	0.03889	0.8247
Kota Depok	128.81976	-1.31610	0.03892	0.8832
Kota Cimahi	128.58999	-1.31372	0.03883	0.8656
Kota Tasikmalaya	128.34288	-1.31116	0.03874	0.8507
Kota Banjar	128.24196	-1.31013	0.03871	0.6673

Berdasarkan hasil yang disajikan pada Tabel 4.1, maka dapat disimpulkan bahwa ada dua wilayah yang memiliki nilai peluang yang besar yaitu Kota Cirebon dan Kabupaten Bandung Barat. Kedua wilayah tersebut memiliki peluang yang besar dalam menekan ABH untuk berada dibawah target 4.8%. Sedangkan wilayah yang memiliki peluang yang cukup kecil adalah Kabupaten Majalengka dan Kabupaten Subang, artinya kedua wilayah tersebut memiliki peluang yang cukup kecil dalam menekan ABH untuk berada dibawah target 4.8%.

Pemilihan Model Regresi Terbaik

Pemilihan model regresi terbaik diperoleh dengan membandingkan nilai AIC pada model Regresi Logistik dengan AIC pada model GWLR. Hasilnya disajikan pada Tabel 3.2 dengan bantuan software *GWR4* adalah sebagai berikut :

Tabel 3.2

Nilai AIC Pada Model Regresi Logistik dan GWLR

Model	AIC
	X_3 - X_4
Regresi Logistik	28.839
GWLR (<i>Bisquare</i> Kernel)	28.824

Pada Tabel 3.2 menunjukkan bahwa nilai AIC model GWLR dengan fungsi pembobot *Bisquare* Kernel tidak terlalu berbeda dibandingkan dengan AIC model regresi logistik. Akan tetapi model GWLR tetap memberikan nilai AIC yang lebih kecil, sehingga dapat dikatakan model GWLR lebih baik dibandingkan model regresi logistik.

D. Kesimpulan

Variabel-variabel yang berpengaruh signifikan pada taraf nyata 10% terhadap model dengan menggunakan model regresi logistik adalah APS (X_3) dan daerah berstatus kota (X_4).

Selanjutnya pengujian dengan metode GWRL pada taraf nyata 10 %, diperoleh hasil yang sama yaitu variabel APS (X_3) dan daerah berstatus kota (X_4) yang berpengaruh signifikan dalam menekan ABH untuk berada di bawah target 4.8% di masing-masing kabupaten/kota Jawa Barat.

Pada model GWLR dan regresi logistik memberikan hasil yang tidak terlalu berbeda berdasarkan kriteria nilai AIC, hal ini menunjukkan bahwa tidak ada pengaruh efek spasial pada pemodelan ABH di Provinsi Jawa Barat. Model GWLR tetap memberikan nilai AIC yang lebih kecil, sehingga dapat dikatakan model GWLR lebih baik dibandingkan model regresi logistik.

Daftar Pustaka

- Anggraini, R. dan Purnadi. 2012. Pemodelan Faktor-Faktor Yang Berpengaruh Terhadap Prevalensi Balita Kurang Gizi Di Provinsi Jawa Timur Dengan Pendekatan *Geographically Weighted Logistic Regression* (GWLR). Jurnal Sains dan Seni ITS Vol. 1, No. 1, http://ejournal.its.ac.id/index.php/sains_seni/issue/view/ISSN%202301-928X, (diakses pada 19 April 2014).
- Atkinson, P. M., S. E. German, D. A. Sear, dan M. J. Clark. 2003. *Exploring The Relations Between Riverbank Erosion and Geomorphological Controls Using Geographically Weighted Logistic Regression*. Ohio : Ohio State University Press.
- Brundson C., A. S. Fotheringham, dan M. E. Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationship*. New York : John Wiley & Sons.
- BPS. 2010. Peraturan Kepala Badan Pusat Statistik No 37 Tahun 2010 Tentang Klasifikasi Perkotaan dan Pedesaan di Indonesia. Badan Pusat Statistika. Jakarta.
- BPS. 2012. Indeks Pembangunan Manusia 2012. Badan Pusat Statistika. Jakarta.
- BPS. 2013. *Persentase Penduduk Miskin*. Badan Pusat Statistik, [Online], <http://sirusa.bps.go.id/index.php?r=indikator/view&id=18> (diakses tanggal 10 Mei 2014).
- Data Statistik Indonesia. 2012. *Angka Partisipasi Sekolah*, [Online], <http://www.datastatistikindonesia.com> (diakses tanggal 17 Mei 2014).
- Data Statistik Indonesia. 2012. *Pengangguran Terbuka* [Online], <http://www.datastatistikindonesia.com> (diakses tanggal 17 Mei 2014).

- Eka Firmansyah, B dan Sutikno. 2011. *Pemodelan dan Pemetaan Angka Buta Huruf Provinsi Jawa Timur dengan Pendekatan Regresi Spasial*, Jurnal Institut Teknologi Sepuluh November, Surabaya, 2011. <http://digilib.its.ac.id/ITS-Undergraduate-3100011043853/16342>, (diakses tanggal 19 April 2014).
- Giovani Elian, D. C., M. Ratna, dan N. Budiantara. 2013. *Pemodelan Faktor-Faktor yang Mempengaruhi Angka Buta Huruf (ABH) di Kabupaten/Kota Jawa Timur dengan Regresi Spline Semiparametrik*. Jurnal Institut Teknologi Sepuluh November, Surabaya, 2013. <http://digilib.its.ac.id/ITS-Undergraduate-13001130003337/26755>, (diakses tanggal 19 April 2014).
- Hosmer, D.W. dan S. Lemeshow. 2000. *Applied Logistic Regression*. New York : John Wiley and Sons.
- Kutner, M. H., Nachtsheim, C.J., Neter, J. dan Li, W. 2004. *Applied Linear Statistical Models*, Fifth Edition, New York : Mc Graw Hill.
- Lailiyah, N. dan Puhadi. 2012. *Pemodelan Faktor-Faktor yang Mempengaruhi Tingkat Buta Huruf Kabupaten/Kota di Jawa Timur dengan Geographically Weighted Ordinal Logistic Regression*, Jurnal Sains dan Seni ITS Vol. 1, No. 1, http://ejournal.its.ac.id/index.php/sains_seni/issue/view/ISSN%202301-928X, (diakses tanggal 19 April 2014).
- Mendikbud 2014. Jumlah Penduduk Buta Huruf, [Online], <http://www.kemdiknas.go.id/kemdikbud/berita/1982>, (diakses tanggal 17 Agustus 2014).