BAB II

TINJAUAN PUSTAKA

2.1 Regresi linier

Analisa Regresi menurut Sudjana dalam (Azizah, 2013) merupakan metode yang digunakan dalam menentukan pola hubungan variabel dependent dengan variabel independent, sehingga didapatkan suatu persamaan pemodelan matematis yang menyatakan hubungan fungsional antara variabel terkait. Variabel dalam analisa regresi dibagi menjadi dua macam antara lain variabel dependent dan variabel independent. Variabel dependent adalah variabel yang dipengaruhi oleh variabel independent, dimana variabel dependent ini muncul akibat adanya variabel independent yang mempengaruhinya. Sedangkan variabel independent adalah variabel yang mempengaruhi prediksi dari variabel dependent dan variabel ini memiliki kedudukan sebagai penjelas dari variabel dependent. Variabel dependent dan independent mengunakan skala interval. Model persamaan umum analisa regresi sebagai berikut:

$$\boldsymbol{y = \beta x + \varepsilon} \tag{1}$$

Menurut Dajan dalam (Syukriyah, 2011) Parameter model regresi di estimasi dengan metode meminimumkan jumlah kuadrat error atau sering dikenal dengan *ordinary least square* (OLS). Dalam OLS errornya diasumsikan identik independen dan berdistribusi normal dengan mean nol dan varians kontans. Dengan taksiran sebagai berikut:

$$\boldsymbol{\beta} = (\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-\mathbf{1}}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{Y} \tag{2}$$

Untuk pengujian model regresi digunakan dua uji yaitu:

a. Uji Kesesuaian Model digunakan untuk menguji kesesuaian model regresi yang dimaksudkan untuk mengetahui pengaruh hubungan antara variabel **X** dan **Y** secara bersamaan (simultan), dalam menguji kesesuaian model digunakan prosedur uji hipotesis sebagai berikut :

$H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0$, (Variabel independent tidak berpengaruh signifikan terhadap variabel dependent secara simultan)

$\mathrm{H}_1 : \beta_p \neq 0$, (Variabel independent berpengaruh signifikan terhadap variabel dependent secara simultan)

Statistik Uji yang digunakan adalah $F_{hitung} = \frac{MSR}{MSE}$. Pengambilan keputusan dengan kriteria tolak H_0 jika $F_{hitung} > F_{a,k,n-k-1}$.

b. Uji Parameter Model pengujian ini digunakan untuk mengetahui variabel prediktor mana saja yang secara signifikan berpengaruh terhadap variabel respon. Dalam melakukan pengujian parameter model digunakan prosedur hipotesis sebagai berikut:

$H_0 : \beta_k = 0$, (Variabel independent tidak berpengaruh signifikan terhadap variabel dependent secara parsial)

$H_1 : \beta_k \neq 0$, (Variabel independent berpengaruh signifikan terhadap variabel dependent secara parsial)

Statistik uji yang digunakan adalah $t_{hit} = \frac{\beta_j}{\sqrt{Se(\beta_j)}}$. Dengan aturan penolakan H_0 jika nilai $|t_{hitung}| > t_{\frac{\alpha}{2},n-k-1}$

2.2 Uji asumsi Residual

Residual menurut Greene dalam (L. Farida, 2010) merupakan selisih antara nilai pengamatan dependent dengan nilai dugaannya atau dapat dikatakan sebagai error (kesalahan) dari pengamatan pada model dengan data keseluruhaannya. Residual dapat dinyatakan dalam fungsi ε. Residual yang baik harus memenuhi asumsi residual yang biasa disebut dengan asumsi klasik, karena asumsi klasik merupakan salah satu syarat yang harus dipenuhi oleh analisa regresi linier sehingga estimasi yang didapatkan unbias atau pemodelan yang dihasilkan dapat dipertanggungjawabkan.

Menurut Islam (2017) untuk mendapatkan model regresi linier berganda dengan estimasi *Linier Square (LS)* yang bersifat *Best Line Unbiased Estimator (BLUE)* yang artinya penduga parameter menghasilkan penduga yang bersifat unbias dan memiliki variansi minimum maka suatu model regresi harus memenuhi beberapa asumsi klasik. Asumsi klasik yang harus dipenuhi adalah residual harus identik (Uji Heterogenitas), residual harus independent (Uji Autokorelasi) dan residual harus berdistribusi normal (Uji Normalitas), akan tetapi berkembangnya zaman asumsi klasik secara umum terbagi menjadi empat asumsi yang harus terpenuhi antara lain asumsi yakni residual harus identik (Uji Heterogenitas), residual harus independent (Uji Autokorelasi) dan residual harus berdistribusi normal (Uji Normalitas) dan Uji Multikolinieritas. Dibawah ini akan diuraikan empat asumsi klasik yang harus terpenuhi oleh model regresi berganda berdasarkan Islam (2017) antara lain:

1. Normalitas

Normalitas merupakan salah satu asumsi klasik yang penting dalam pemodelan regresi linier. Normalitas ini digunakan untuk melihat sebaran data residual berdistribusi normal atau tidak. Uji yang secara umum digunakan untuk menguji normalitas antara lain uji histogram, uji normal P-Plot, uji Chi-Square, Skewness dan Kurtosis atau Kolmorgorov Smirnov. Data yang berdistribusi normal mempunyai ciri-ciri sebagai berikut:

a. Kurva distribusi normal berbentuk simetrik yang membentuk kurva seperti lonceng dengan frekuensi tertinggi berada di tengah-tengah (mean) kurva sejajar dan sisi kanan kirinya tepat.
b. Kurva distribusi normal berbentuk simetris sehingga nilai-nilai diatas rata-rata (mean) akan cocok dengan frekuensi nilai dibawah rata-rata (mean) yang merupakan frekuensi total semua nilai dalam populasi.
c. Kurva normal bergantung pada mean dan nilai simpangan baku (standar deviasi), sehingga bentuk kurva normal akan berbeda-beda sesuai dengan nilai mean dan simpangan baku (standar deviasi).

Suatu residual berdistribusi normal harus memenuhi kriteria yang sesuai dengan parameter yang digunakan. Pada tabel 2.1 disajikan tabel kriteria residual berdistribusi normalitas.

Tabel 2.1 Kriteria residual berdistribusi normal

Parameter	**Kriteria Normal**
Koefisien Varian	Nilai Koefisien varian < 30%
Rasio Skewness	Nilai Rasio Skewness berada pada -2 s/d 2
Rasio Kurtosis	Rasio kurtosis berada pada -2 s/d 2
Histogram	Bentuk simetris, berada di tengah, tidak tinggi maupun rendah
Box Plot	Simetris, median tepat di tengah dan tidak ada nilai ekstrim
Normal Q-Q Plot	Data menyebar sekitar garis
Detrended Q-Q Plot	Data menyebar sekitar garis pada nilai 0

Selain kriteria yang ditampilkan diatas pengujian normalitas mengunakan prosedur hipotesis yang disajikan sebagai berikut:

$H_0 : F(x) = F_0(x)$ untuk semua nilai x (Residual berdistribusi normal)

$H_1 : F(x) \neq F_0(x)$ untuk minimal satu x (Residual tidak berdistribusi normal)

Dengan aturan penolakan apabila nilai p-value lebih kecil dari α maka H0 ditolak dan berlaku H1 dan apabila p-value lebih besar dari α maka H0 diterima. Jadi residual berdistribusi normal apabila nilai p-value lebih besar dari α.

2. Autokorelasi

Autokorelasi digunakan untuk melihat apakah variabel residual terjadi korelasi antara variabel residual satu dengan variabel residual yang lain, dimana asumsi least square klasik residual bersifat independent (saling bebas). Autokorelasi dalam konsep regresi menurut Setiawan dan Endah dalam (Azizah, 2013) merupakan komponen residual yang berkorelasi berdasarkan urutan waktu, ruang dan pada dirinya sendiri. Apabila tidak terjadi pelanggaran asumsi autokorelasi maka kovarian antara residual satu (ε_i) dengan residual yang lain (ε_i) bernilai sama dengan nol yang artinya bahwa komponen residual yang berkaitan dengan data pengamatan ke-i tidak dipengaruhi oleh residual yang berkaitan dengan data pengamatan ke-j. Secara sistematis ditulis dengan persamaan sebagai berikut:

$$Cov(\varepsilon_i \varepsilon_j) = E\,\{[\varepsilon_i - E(\varepsilon_i)][\varepsilon_j - E(\varepsilon_j)]\} \quad (3)$$

$$= E(\varepsilon_i \varepsilon_j) = 0; \quad i \neq j \quad (4)$$

Alat yang biasa digunakan untuk mendeteksi adanya autokorelasi yakni mengunakan uji Durbin Watson, dimana pada asumsi autokorelasi digunakan prosedur pengujian hipotesis sebagai berikut:

$H_0 :\ \rho = 0$ (Residual tidak terjadi pelangaran Autokorelasi)

$H_0 :\ \rho \neq 0$ (Residual terjadi pelangaran Autokorelasi)

Aturan penolakan hipotesis mengunakan nilai α yakni apabila nilai p-value (probabilitas) lebih besar dari α maka tolak H1 dan berlaku H0, sedangkan nilai p-value (probabilitas) lebih kecil dari α maka tolak H0 dan berlaku H1. Jadi terjadinya pelanggaran asumsi klasik pada Autokorelasi apabila nilai p-value (probabilitas) lebih kecil dari α . Apabila residual mengandung korelasi maka terdapat 2 cara menangulanginya yang pertama melakukan estimasi metode *generalized least square* dengan melakukan respesifikasi model dengan memasukan komponen lag variabel independent maupun dependent dalam model mengunakan pendekatan estimasi Newey-West yang bersifat *heterocedasticity and autocorrelation consistent (HAC).* Cara kedua mentransformasi data atau mengubah model regresi kedalam bentuk persamaan umum yang berbeda *(generalized difference equation).* Cara ketiga memasukan variabel lag dari variabel dependent menjadi salah satu variabel independent, sehingga data pengamatannnya berkurang 1.

3. Heteroskedastisitas

Sebelum masuk pada uji Heterokesdastisitas alangkah baiknya untuk mengetahui Homoskedastisitas, karena residual yang baik harus identik. Homoskesdastisitas adalah gambaran data dimana varian batas kesalahan yang identik diluar jangkauan nilai-nilai variabel bebas tertentu. Ketika batas kesalahan memiliki varians yang semakin besar maka di indikasikan bersifat heterokesdastisitas, atau dapat dikatakan homoskesdastisitas merupakan asumsi dimana variabel dependent menunjukkan tingkatan varian yang sama untuk semua variabel independent.

Menurut Azizah (2013) Uji Heterokesdastisitas digunakan untuk melihat apakah terdapat ketidaksamaan variansi pada residual pengamatan satu ke pengamatan yang lain. Cara mendeteksi uji heterokesdastisitas dengan 3 cara yakni mengunakan uji glejser, mengunakan visualisasi grafis dan mengunakan cara formal dengan prosedur hipotesis. Uji glejser ini mengunakan hasil residual kuadrat terhadap prediktornya. Sedangkan cara visualisasi yakni dengan cara melihat pada pola nonacak dari plot residual atau residual kuadratis terhadap variabel independent terhadap nilai variabel dependen dengan catatan model telah di estimasi. Cara ketiga yakni secara formal mengunakan prosedur hipotesis sebagai berikut:

H_0 : $\beta_j = 0$ (Residual memiliki nilai varians yang identik / Homokesdastisitas)

H_0 : $\beta_j \neq 0$ (Residual tidak memiliki nilai varians yang identik / Hetokesdastisitas)

Aturan penolakan hipotesis mengunakan nilai α yakni apabila nilai p-value (probabilitas) lebih besar dari α maka tolak H1 dan berlaku H0, sedangkan nilai p-value (probabilitas) lebih kecil dari α maka tolak H0 dan berlaku H1. Jadi terjadinya pelanggaran asumsi klasik pada Heterokesdatisitas apabila nilai p-value (probabilitas) lebih kecil dari α .

Menurut Azizah (2013) dampak terjadinya pelangaran asumsi heteroskesdastisitas antara lain:

a. Statistik uji t (parsial) pada pengujian parameter regresi menjadi tidak valid.
b. Parameter regresi memiliki selang kepercayaan yang melebar, hal ini akan berakibat pada hasil perkiraan yang diperoleh tidak dipercaya.

Menurut Gujarati dalam (Azizah, 2013) dampak yang ditimbulkan ketika terjadi pelanggaran asumsi heterokesdastisitas dikarenakan ganguan pada variansi yang melebar yakni estimator kuadrat terkecil (OLS) tetap linier, tak bias, tidak memiliki varians minimum yang beakibat tidak efisiennya varians dan hal ini berlaku pada sampel besar, pengujian tes hipotesis mengunakan uji F (simultan) dan t (parsial) tidak meyakinkan, rumus penaksir varians kuadrat terkecil (OLS) akan bias,dimana bias positif terjadi ketika taksiran varian pada OLS terlalu besar dan bias negatif apabila taksiran varians OLS terlalu kecil.

Cara mengatasi pelanggaran Heterokedastisitas menurut Setiawan dan Endah dalam (Azizah, 2013) antara lain:

a. Mentransformasikan variabel dependen maupun variabel independent, dimana fungsi yang digunakan dalam transformasi ini yakni ln, log,√ (akar), sinus, kosinus, Box-Cox, $\frac{1}{Y}$, $\frac{1}{X}$, dan lain sebagainya.
b. Mengatasi pelanggaran asumsi heterokesdastisitas yakni mengunakan metode kuadrat terkecil tertimbang, dimana metode ini mengunakan matriks pembobot berupa matriks diagonal. Dalam melakukan perkiraan pembobot yang tidak diketahui maka harus memenuhi asumsi perkiraan pembobot yakni variansi residual proporsional ke X^2 sehingga digunakan transformasi $\frac{1}{Y}$ dan variansi residual proporsional ke X sehingga digunakan transformasi $\frac{1}{\sqrt{X}}$.

4. Multikolinieritas

Pelanggaran asumsi multikolinieritas terjadi ketika terjadi adanya hubungan linier antar variabel independent, apabila terjadi keeratan hubungan yang kuat antar variabel independet maka akan terjadi gangguan pada hubungan dependent dengan independen, hal ini berakibat pada estimasi koefisien menjadi tidak valid. Menurut Syukriyah (2011) dalam mendeteksi pelanggaran multikolinieritas digunakan nilai *tolerance (TOL)* dan *varians infloating factor*

(VIF). Definisi persamaan *tolerance (TOL)* dan *varians infloating factor (VIF)* disajikan sebagai berikut:

$$TOL = 1 - R^2 \tag{5}$$

$$VIF = \frac{1}{TOL} \tag{6}$$

Dengan aturan terjadinya pelanggaran Multikolinieritas apabila nilai *tolerance (TOL)* lebih kecil dari 0,1 dan *varians infloating factor (VIF)* lebih besar dari 10. Selain mengunakan *tolerance (TOL)* dan *varians infloating factor (VIF)* untuk mendeteksi pelangaran multikolinieritas, ada juga yang mengunakan prosedur hipotesis yang dinyatakan sebagai berikut:

H_0 : Tidak terjadi pelanggaran multikolinieritas

H_1 : Terjadi pelanggaran multikolinieritas

Aturan penolakan hipotesis mengunakan nilai α yakni apabila nilai p-value (probabilitas) lebih besar dari α maka tolak H1 dan berlaku H0, sedangkan nilai p-value (probabilitas) lebih kecil dari α maka tolak H0 dan berlaku H1. Jadi terjadinya pelanggaran asumsi klasik pada Multikolinieritas apabila nilai p-value (probabilitas) lebih kecil dari α.

2.3 Heterogenitas Spasial

Pengujian heterogenitas spasial menurut Purhadi dan Yasin dalam (Widiyanti dkk., 2014) merupakan efek yang terjadi karena adanya perbedaan karakteristik suatu lokasi satu dengan lokasi pengamatan lainnya (efek wilayah yang random) dimana pada pengujian heterogenitas spasial ini sangat penting karena apabila asumsi heterogenitas spasial tidak terpenuhi akan berpengaruh pada hasil yang diperoleh tidak efisien dalam estimasi dan kesimpulan yang didapatkan kurang sahih. Asumsi heterogenitas spasial dilakukan dengan mengunakan statistik uji *Breush-Pagan Test.* Dengan hipotesis yang digunakan dalam uji *Breush-Pagan Test* sebagai berikut:

$H_0 : \sigma_i^2 = 0$ (Data tidak terjadi Heterogenitas Spasial)

$H_1 : \sigma_i^2 \neq 0$ (Data terjadi Heterogenitas Spasial)

Dengan aturan penolakan H0 apabila nilai *Breush-Pagan Test (BP)* < X^2 maka tolak H0 dan berlaku H1, sedangkan nilai *Breush-Pagan Test (BP)* > X^2maka terima H0. Sehingga terpenuhi Heterogenitas spasial apabila *Breush-Pagan Test (BP)* < X^2.

2.4 *Geographically Weighted Regression (GWR)*

Geographically Weighted Regression (GWR) menurut Mei et al. dalam (Yasin, 2013) adalah pengembangan model regresi global yang berdasarkan regresi non parametrik, dimana setiap lokasi pengamatan mempunyai nilai parameter regresi yang berbeda-beda sehingga model ini menghitung parameter setiap lokasi pengamatannya. Model *Geographically Weighted Regression (GWR)* mempunyai asumsi yang harus terpenuhi yakni asumsi residual harus berdistribusi normal dengan mean sama dengan nol dan varians σ^2. Secara umum persamaan *Geographically Weighted Regression (GWR)* dapat ditulis sebagai berikut:

$$\boldsymbol{y_i} = \boldsymbol{\beta_0}(\boldsymbol{u_i}, \boldsymbol{v_i}) + \sum_{k=1}^{p} \boldsymbol{\beta_k}(\boldsymbol{u_i}, \boldsymbol{v_i})\boldsymbol{x_{ik}} + \boldsymbol{\varepsilon_i} \quad (7)$$

Estimasi parameter model *Geographically Weighted Regression (GWR)* mengunakan pembobot yang berbeda untuk setiap lokasi pengamatan, metode yang memberi pembobot yang berbeda disetiap daerah pengamatan disebut *Weighted Least Square (WLS).* Secara umum persamaan pembobot disajikan sebagai berikut:

$$\boldsymbol{\beta}(\boldsymbol{u}_i, \boldsymbol{v}_i) = (\boldsymbol{X}^T \boldsymbol{W}(\boldsymbol{u}_i, \boldsymbol{v}_i)\boldsymbol{X}^{-1}\boldsymbol{X}^T \boldsymbol{W}(\boldsymbol{u}_i, \boldsymbol{v}_i)\boldsymbol{y} \quad (8)$$

Dengan $\boldsymbol{x}^T = (\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \dots, \boldsymbol{x}_{ip})$ merupakan elemen baris ke-I dari matriks X, sehingga nilai prediksi untuk y pada $(u_i v_i)$ disajikan sebagai berikut:

$$\boldsymbol{y}_i = \boldsymbol{x}_i^T \boldsymbol{\beta}(\boldsymbol{u}_i, \boldsymbol{v}_i) = \boldsymbol{x}_i^T (\boldsymbol{X}^T \boldsymbol{W}\,(\boldsymbol{u}_i, \boldsymbol{v}_i)X)^{-1} \boldsymbol{X}^T \boldsymbol{W}(\boldsymbol{u}_i, \boldsymbol{v}_i)\boldsymbol{y} \quad (9)$$

Menurut Lesage **W(i)** merupakan matriks pembobot spasial yang diperoleh dari fungsi kernel seperti fungsi jarak Gaussian (Gaussian Function), fungsi Exponential berdasarkan jarak Euclid antar wilayah, sedangkan menurut Chasco matrik pembobot spasial yang diperoleh dari fungsi Bisquaredan fungsi kernel tricube berdasarkan jarak Euclid antar wilayah. (Yasin et al., 2018)

Menurut Fotheringham dalam (Huang et al., 2010) matriks pembobot *Geographically Weighted Regression (GWR)* mewakili karakteristik setiap lokasi pengamatan yang berbeda, semakin dekat dengan lokasi pengamatan maka semakin besar pembobotnya. Jadi setiap estimasi titik mempunyai matrik pembobot yang berbeda. Terdapat dua pembobot yang digunakan yakni fixed kernel dan adaptive kernel. Pada fixed kernel memiliki jarak yang tetap akan tetapi jumlah tetangga bervariasi, sedangkan pada kernel adaptive memiliki jarak yang bervariasi akan tetapi jumlah tetangga tetap. Secara umum fungsi yang banyak digunakan adalah fungsi kernel peluruhan jarak *distance Gaussian* dengan persamaannya sebagai berikut:

$$W_y = \exp(\frac{d_y^2}{h^2}) \tag{10}$$

Dimana h adalah parameter non-negatif yang dikenal sebagai bandwith yang menghasilkan peluruhan pengaruh jarak dan ukuran jarak antara lokasi i dan j (d_{ij}). $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ dan $(\boldsymbol{x}_j, \boldsymbol{y}_j)$ merupakan titik koordinat yang digunakan dalam *Geographically Weighted Regression (GWR)* dan jarak yang biasa digunakan dalam metode *Geographically Weighted Regression (GWR)* adalah jarak Euclidean, yang disajikan pada persamaan sebagai berikut:

$$d_{ij} = \sqrt{(\boldsymbol{x}_i - \boldsymbol{x}_j)^2 + (\boldsymbol{y}_i - \boldsymbol{y}_j)^2} \tag{11}$$

Menurut Fotheringham dalam (Widiyanti et al., 2014) salah satu metode yang banyak digunakan dalam penentuan bandwith optimum yakni mengunakan *Cross Validation (CV).* Secara matematis dirumuskan pada persamaan tersebut:

$$CV(h) = \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq 1}(h)\right)^2 \tag{12}$$

Dimana $\hat{y}_{\neq 1}(h)$ adalah penduga $\boldsymbol{y}_i$ dan lokasi pengamatan (u_i, v_i) dihilangkan dalam proses pendugaan. Menurut Fotheringham dalam (Huang et al., 2010) parameter h dapat secara otomatis mengoptimalkan dengan cara melihat nilai minimal pada statistic *goodness of fit* atau nilai AIC.

2.5 *Geographically and Temporally Weighted Regression (GTWR)*

Menurut Wang dalam (Huang et al., 2010) dalam praktik pemodelan *Geographically Weighted Regression (GWR)* menjelaskan nonstasioner spasial pada estimasi dengan cara membangun pembobot spasial berdasarkan jarak antara titik estimasi i dan pengamatan lain. Akan tetapi secara umum, variabel waktu diakomodasi secara terpisah dengan menyesuaikan pengamatan harga jual dengan waktu, sehingga sering menggunakan beberapa bentuk nilai kini yang disesuaikan atau perhitungan nilai masa depan.

Menurut Wang dalam (Widiyanti et al., 2014) *Geographically and Temporally Weighted Regression (GTWR)* muncul untuk menangani ketidak stasionernya suatu data pada sisi spasial maupun temporal dalam metode *Geographically Weighted Regression (GWR)*, dimana metode *Geographically and Temporally Weighted Regression (GTWR)* mengabungkan informasi lokasi dan waktu pada matriks pembobot sehingga dapat mengidentifikasi adanya heterogenitas spasial dan temporal. Secara umum persamaan *Geographically and Temporally Weighted Regression (GTWR)* disajikan sebagai berikut:

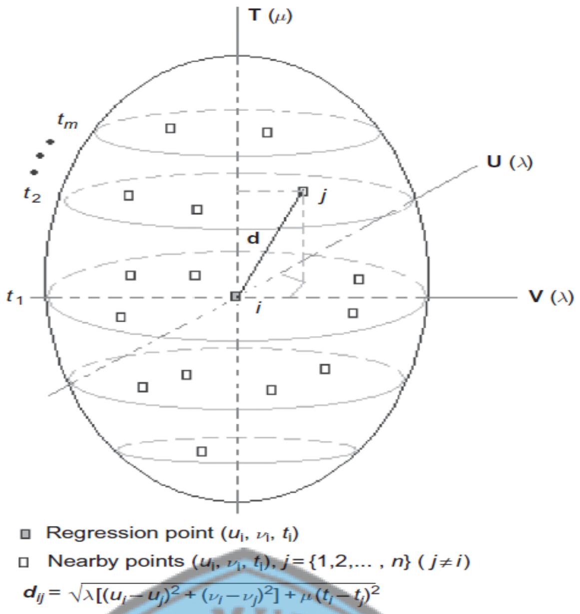
$$\boldsymbol{y_i = \beta_0(u_i, v_i, t_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i, t_t) x_{ik} + \varepsilon_i} \quad (13)$$

Menurut Huang et al. (2010) permasalahan yang dihadapi disini yakni menentukan niai taksiran dari $\boldsymbol{\beta_k(u_i, v_i, t_t)}$ untuk setiap variabel k dan setiap lokasi ruang waktu ke-i, sehingga didapatkan persamaan estimasi $\boldsymbol{\beta_k(u_i, v_i, t_t)}$ sebagai berikut:

$$\boldsymbol{\beta(u_i, v_i, t_i) = (X^T W(u_i, v_i, t_i) X)^{-1} X^T W(u_i, v_i, t_i) y} \quad (14)$$

Dimana $\boldsymbol{W}(\boldsymbol{u_i}, \boldsymbol{v_i}, \boldsymbol{t_i}) = diag(\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{in})$ dan n merupakan banyaknya pengamatan, dimana elemen diagonal $\alpha_{ij}(1 \leq j \leq n)$ yang merupakan fungsi jarak spasial-temporal pada titik pengamatan(u_i, v_i, t_i) yang diasumsikan bahwa kedekatan titik pengamatan data terhadap titik i pada koordinat spasial-temporal mempunyai pengaruh yang lebih besar dalam estimasi parameter $\boldsymbol{\beta}(\boldsymbol{u_i}, \boldsymbol{v_i}, \boldsymbol{t_i})$ daripada data yang mempunyai lokasi yang jauh dari titik pengamatan ke-i.

Menurut Huang et al. (2010) sebelum menghitung fungsi jarak spasial-temporal, alangkah lebih baiknya jika memahami beberapa gagasan yang menjadi dasar dalam pengukuran kedekatan, misalkan data yang diamati terletak pada tiga dimensi pada sistem koordinat spasial-temporal dan memiliki kedekatan dengan titik i yang digambarkan mengunakan bola jari-jari tertentu, katakanlah r berada di sekitar titik regresi i sehingga tolak ukur yang digunakan pada metode ini adalah metode Ordinary least square (OLS) hanya dalam lingkup pengamatan ini. $k(u_i, v_i, t_i)$ yang diperoleh dianggap sebagai perkiraan asosiasi antar variabel di dalamnya dan sekitar titik i. Penjelasan tersebut disajikan pada ilustrasi dibawah ini:



Gambar 2.1 ilustrasi jarak spasial-temporal

Fungsi jarak spasial-temporal terdiri dari gabungan fungsi jarak spasial (d^S) dan fungsi jarak temporal (d^T) sehingga didapatkan persamaan sebagai berikut:

$$d^{ST} = \lambda(d^S)^2 + \mu(d^T)^2 \tag{15}$$

Dimana λ dan μ menyatakan faktor skala penyeimbang efek yang berbeda dalam melakukan penghitungan jarak spasial-temporal, sehingga didapatkan estimasi jarak euclidean sebagai berikut:

$$\left(d_{ij}^{ST}\right)^2 = \lambda\,\{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2 + \mu\,\left(t_i - t_j\right)^2 \tag{16}$$

Dari persamanaan diatas diperoleh estimasi sebagai berikut:

$$\alpha_{ij} = \exp\{-(\frac{\lambda\,[\left(u_i-u_j\right)^2+\left(v_i-v_j\right)^2+\mu\left(t_i-t_j\right)^2}{h_{ST}^2})\,\} \tag{17}$$

$$\alpha_{ij} = \exp\{-(\frac{\left(u_i-u_j\right)^2+\left(v_i-v_j\right)^2}{h_{ST}^2} + \frac{\left(t_i-t_j\right)^2}{h_{ST}^2})\,\} \tag{18}$$

$$= \exp\{-(\frac{(d_{ij}^S)^2}{h_S^2} + \frac{(d_{ij}^T)^2}{h_T^2})\} \quad (19)$$

$$= \exp\{-(\frac{(d_{ij}^S)^2}{h_S^2}\} x \exp\{- \frac{(d_{ij}^T)^2}{h_T^2}\} \quad (20)$$

$$= \alpha_{ij}^S \ x \ \alpha_{ij}^T \quad (21)$$

Dimana h_{ST}^2 merupakan parameter dari bandwith spasial-temporal, h_S^2 sebagai parameter bandwith spasial dan h_T^2 sebagai parameter bandwith temporal. Apabila τ di misalkan sebagai parameter rasio dari $\frac{\mu}{\lambda}$ dengan $\lambda \neq 0$ maka diperoleh persamaan:

$$\frac{(d_{ij}^{ST})^2}{\lambda} = (u_i - u_j)^2 + (v_i - v_j)^2 + \tau(t_i - t_j)^2 \quad (22)$$

Menurut Huang et al. (2010) Pengunaan parameter τ bertujuan untuk memperbesar atau memperkecil efek jarak temporal terhadap jarak spasial. Menurut Laksana (2018) Parameter τ didapatkan melalui metode optimasi koefisien determinasi (R^2) secara iteratif. Sehingga dihasilkan estimasi parameter τ yang menghasilkan (R^2) yang maksimum. Dari estimasi parameter τ yang menghasilkan (R^2) yang maksimum sehingga didapatkan estimasi parameter μ dan λ.

2.5.1 Algoritma Metode Iteratif Parameter τ

Langkah analisa untuk mendapatkan parameter τ melalui metode iteratif berdasarkan penelitian Laksana (2018) adalah sebagai berikut:

1. Menentukan nilai awal τ_0 = 0.005

2. Mendapatkan nilai jarak *Euclidean* yang diperoleh dari persamaan (22)
3. Menghitung nilai fungsi jarak pembobot yang diperoleh dari persamaan sebagai berikut.

$$K_{ij} = \exp(-\frac{(d_{ij}^{st})^{\wedge}2\ /\lambda}{h_s^2}) \quad (23)$$

4. Mendapatkan nilai estimasi β berdasarkan persamaan 14 dengan menggunakan matriks pembobot yang diperoleh dari langkah (3)
5. Melakukan perhitungan *Sum Square Error* (SSE) dan *Sum Square Total* (SST)
6. Sehingga diperoleh nilai koefisien determinasi (R^2)
7. Melakukan iterasi langkah 2 sampai 6 dengan nilai $\tau s + 1 = s \times \tau s$ dimana $s = 0,1,2,\ldots,n$
8. Mendapatkan nilai parameter τ dengan cara memilih estimasi R^2 optimum. R^2 optimum ditentukan dengan nilai R^2 yang mulai konstan.

2.5.2 Algoritma Metode Iteratif Estimasi Parameter μ dan λ

Langkah analisa untuk mendapatkan parameter μ dan λ melalui metode iteratif berdasarkan penelitian Laksana (2018) sebagai berikut:

1. Menentukan nilai awal μ0 = 0.03 dan λ0 = 0.012 kemudian dikalikan pembanding yang didapatkan dari parameter τ .
2. Mendapatkan nilai jarak *Euclidean* yang diperoleh dari persamaan 16.
3. Menghitung fungsi jarak pembobot yang diperoleh dari persamaan berikut:

$$K_{ij} = \exp(-\frac{(d_{ij}^{st})^{\wedge}2}{h_s^2}) \quad (24)$$

4. Mendapatkan nilai estimasi β berdasarkan persamaan 14 dengan menggunakan matriks pembobot yang diperoleh dari langkah 3
5. Melakukan perhitungan *Sum Square Error* (SSE) dan *Sum Square Total* (SST).
6. Sehingga diperoleh nilai koefisien determinasi (R^2)
7. Melakukan iterasi langkah 2 sampai 6 dengan nilai $\mu s + 1 = s \times \mu s$ dan $\lambda s + 1 = s \times \lambda s$ dimana $s = 0,1,2, \dots, n$
8. Mendapatkan nilai parameter μ dan λ dengan melakukan pemilihan estimasi R^2 optimum. R^2 optimum ditentukan dengan nilai R^2 yang mulai konstan.

2.6 Pengujian Hipotesis *Geographically and Temporally Weighted Regression (GTWR)*

Menurut Leung et al. dalam (Widiyanti et al., 2014) pengujian hipotesis *Geographically and Temporally Weighted Regression (GTWR)* terdiri dari pengujian kesesuaian model dan pengujian parameter model. Pada pengujian hipotesis pertama yakni pengujian kesesuaian model *Geographically and Temporally Weighted Regression (GTWR)* tidak berbeda dengan pengujian kesesuaian model *Geographically Weighted Regression (GWR)* yang dilakukan dengan prosedur hipotesis sebagai berikut:

$$H_0: \beta_k(u_i, v_i, t_i) = \beta_k \ untuk\ setiap\ pengamatan\ k, i = 0,1,2, \dots, p.$$

(Tidak adanya perbedaan yang signifikan antara model regresi global dengan GTWR)

H_1: $paling\ sedikit\ satu\ \beta_k(u_i, v_i, t_i) \neq \beta_k$. (Adanya perbedaan yang signifikan antara model regresi global dengan GTWR)

Statistik uji yang digunakan dalam penentuan hipotesis ini mengunakan *Residual Sum of Square (RSS)* dengan persamaannya sebagai berikut:

$$F_1 = \frac{RSS(H_1)/(\frac{\delta_1^2}{\delta_2^1})}{\frac{RSS(H_0)}{n-p-1}} \tag{25}$$

Dimana $\left(\frac{\delta_1^2}{\delta_2^1}\right)$ sebagai derajat bebas 1 (df_1) dan (n-p-1) sebagai derajat bebas 2 (df_2) sehingga aturan penolakan H_0 apabila $F_1 < F_{1-\alpha, df_1, df_2}$. Menurut Mei et al. dalam (Widiyanti et al., 2014) apabila pengujian kesesuaian model didapatkan hasil model GTWR berbeda dengan model regresi global, sehingga di uji lanjut secara parsial untuk mengetahui apakah ada pengaruh yang signifikan variabel x_k antara lokasi satu dengan yang lain, dengan prosedur hipotesisnya sebagai berikut:

H_0: $\beta_1(u_1, v_1, t_1) = \beta_2\ (u_2, v_2, t_2) = \cdots = \beta_k\ (u_i, v_i, t_i)$

(Tidak adanya perbedaan pengaruh yang signifikan antara variabel x_k dengan lokasi dan waktu)

H_1: $paling\ sedikit\ satu\ \beta_k(u_i, v_i, t_i)\ yang\ berbeda$.

(Adanya perbedaan pengaruh yang signifikan antara variabel x_i dengan lokasi maupun waktu)

Sebelum melakukan pengujian hipotesis di atas maka harus menentukan terlebih dahulu varians $\beta_k(u_i, v_i, t_i)$ yang di notasikan dengan:

$$V_k^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\beta(u_i, v_i, t_i) - \frac{1}{n}\sum_{i-1}^{n}\beta(u_i, v_i, t_i)\right)^2 \quad (26)$$

$$\sum_{i-1}^{n}\beta_k^T\left[\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}\right]\beta_k \quad (27)$$

Dengan $\boldsymbol{\beta}(\boldsymbol{u_i}, \boldsymbol{v_i}, \boldsymbol{t_i}) = \begin{bmatrix} \boldsymbol{\beta_k}(\boldsymbol{u_1}, \boldsymbol{v_1}, \boldsymbol{t_1}) \\ \boldsymbol{\beta_k}(\boldsymbol{u_2}, \boldsymbol{v_2}, \boldsymbol{t_2}) \\ . \\ . \\ . \\ \boldsymbol{\beta_k}(\boldsymbol{u_i}, \boldsymbol{v_i}, \boldsymbol{t_i}) \end{bmatrix}$ dan **I** adalah matriks identitas berukuran nxn dan **J** merupakan matris berukuran nxn yang semua elemennya adalah 1. Dengan statistik uji yang digunakan:

$$F_3 = \frac{\frac{V_k^2}{tr(\frac{1}{n}\beta_k^T[I - \frac{1}{n}J]\beta_k)}}{RSS\ (H_1)/\delta_1} \quad (28)$$

Dengan $\boldsymbol{\beta_k} = \begin{bmatrix} \boldsymbol{e_k^T}[X^T W(\boldsymbol{u_1}, \boldsymbol{v_1}, \boldsymbol{t_1})X]^{-1} X^T W(\boldsymbol{u_1}, \boldsymbol{v_1}, \boldsymbol{t_1}) \\ \boldsymbol{e_k^T}[X^T W(\boldsymbol{u_2}, \boldsymbol{v_2}, \boldsymbol{t_2})X]^{-1} X^T W(\boldsymbol{u_2}, \boldsymbol{v_2}, \boldsymbol{t_2}) \\ . \\ . \\ . \\ \boldsymbol{e_k^T}[\boldsymbol{X^T W}(\boldsymbol{u_n}, \boldsymbol{v_n}, \boldsymbol{t_n})\boldsymbol{X}]^{-1} \boldsymbol{X^T W}(\boldsymbol{u_n}, \boldsymbol{v_n}, \boldsymbol{t_n}) \end{bmatrix}$

Dimana $\mathbf{e_k}$ adalah vektor korom berukuran (p+1) yang bernilai satu untuk elemen ke-k dan nol untuk lainnya. Statistik uji untuk $\left(\frac{\gamma_1^2}{\gamma_2^1}\right)$ sebagai derajat bebas 1 (df_1) dan $\left(\frac{\delta_1^2}{\delta_2^1}\right)$ sebagai derajat bebas 2 (df_2) sehingga aturan penolakan H_0 apabila $F_3 \geq F_{\alpha, df_1, df_2}$ leung et al. dalam (Widiyanti dkk., 2014).

2.7 Pemilihan Model Terbaik

Pemilihan model terbaik menurut Gujarati dalam (Widiyanti dkk., 2014) yakni metode MSE dan Koefisien Determinasi, dimana pada penelitian ini akan ditambahkan metode dalam pemiliha metode terbaik yakni mengunakan nilai AIC yang didefinisikan sebagai berikut:

a. *Mean Square Error (MSE)*

Mean Square Error (MSE) merupakan salah satu metode dalam pemilihan metode terbaik dengan cara mengkuadratkan masing- masing kesalahan *(error)* dan dijumlahkan serta dibagi dengan jumlah observasi seperti yang disajikan pada estimasi dibawah ini:

$$MSE = \frac{1}{m\,x\,n}\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}[f(i,j) - g(i,j)]^2 \qquad (29)$$

Dengan pemilihan metode terbaik berdasarkan nilai MSE yakni mengunakan hasil nilai MSE terkecil.

b. Koefisien Determinasi (R^2)

Nilai koefisien determinasi ini merupakan salah satuh metode yang menunjukan proporsi atau persentasi variansi total dalam menentukan variabel dependent yang dijelaskan oleh variabel independent atau dapat dikatakan seberapa besar variabel independent mempengaruhi variabel dependent, dengan rumus penghitungan koefisien determinasi (R^2) sebagai berikut:

$$R^2 = 1 - \frac{SSE}{SST} \qquad (30)$$

c. Akaike Information Criterion (AIC)

Akaike Information Criterion (AIC) adalah Salah satu kriteria untuk pemilihan model terbaik berdasarkan nilai yang terkecil. Rumus penghitungan nilai Akaike Information Criterion (AIC) sebagai berikut:

$$AIC = 2n\ln(\sigma^2) + n\ln(2\pi) + n + tr(G) \tag{31}$$

Dengan pemilihan metode terbaik berdasarkan nilai AIC yakni mengunakan hasil nilai AIC terkecil.

2.8 Indeks Pembangunan Manusia (IPM)

Indeks Pembangunan Manusia (IPM) merupakan salah satu indikator pengukuran dari perbandingan harapan hidup, melek huruf, pendidikan dan standar hidup yang ditetapkan oleh seluruh negara di dunia, yang mengklasifikasikan sebuah negara tersebut tergolong ke dalam negara maju, berkembang atau terbelakang sehingga dapat mengukur pengaruh pada kebijakan ekonomi terhadap kualitas hidup. (Heriyanto, 2015).

Konsep Indeks Pembangunan Manusia (IPM) untuk mengukur pencapaian rata-rata kemajuan sebuah negara dalam 3 indikator utama penyusun Indeks Pembangunan Manusia (IPM) yang dijelaskan sebagai berikut:

a. Kehidupan dengan kondisi hidup yang sehat dan mempunyai umurpanjang yang diukur mengunakan angka harapan hidup manusia.

b. Pengetahuan yang diukur dengan angka tingkat baca tulis pada orang dewasa dan kombinasi pendidikan dasar, menengah atas bobot satu per tiga hal ini diukur mengunakan nilai angka melek huruf dan harapan lama sekolah.
c. Logaritma natural dari produk domestik regional bruto per kapita dalam paritasi daya beli menjadi ukuran dalam menentukan standar kehidupan yang layak.

2.9 Tingkat Pengangguran Terbuka (TPT)

Menurut BPS dalam (Utama, 2015) Tingkat Pengangguran Terbuka adalah indikator utama yang digunakan untuk mengukur angka penganguran dalam angkatan kerja.Tingkat Pengangguran Terbuka (TPT) adalah persentase banyaknya pengangguran terhadap banyaknya angkatan tenaga kerja.

Menurut Todaro dalam (Putra, 2015) mengatakan bahwa pembangunan manusia merupakan tujuan pembangunan yang memainkan peranan dalam membentuk kemampuan suatu negara menyerap teknologi modern dalam menciptakan kesempatan kerja dan mengurangi jumlah pengangguran. Apabila permasalahan pengangguran dapat teratasi maka akan berakibat pada efek pendapatan yang tinggi sehingga akan berpengaruh terhadap peningkatan pembangunan manusia melalui peningkatan pengeluaran rumah tangga yang dibelanjakan dalam memenuhi makanan bergizi dan pendidikan yang tinggi. Jadi pengurangan pengangguran dapat dilihat dari nilai Indeks Pembangunan Manusia (IPM).

2.10 Produk Domestik Regional Bruto (PDRB)

Menurut Bhakti et al. (2018) pertumbuhan ekonomi merupakan salah satu indikator dalam mengetahui kinerja perekonomian regional (daerah), dimana pertumbuhan ekonomi adalah kenaikan outup agregat (keseluruhan barang dan jasa yang diperoleh dari kegiatan perekonomian) atau Produk Domestik Regional Bruto (PDRB). Sedangkan menurut Mirza (2012) Pertumbuhan ekonomi adalah persentase nilai yang dilihat dari nilai Produk Domestik Regional Bruto (PDRB) yang dijadikan sebagai tolak ukur dalam meningkatkan perekonomian suatu negara.

Menurut Putra (2015) dalam pembangunan ekonomi atau pertumbuhan ekonomi merupakan syarat bagi tercapainya pembangunan manusia karena dengan pembangunan ekonomi yang meningkat maka akan meningkatkan produktivitas dan peningkatan pendapatan melalui penciptaan kesempatan kerja. Tingkat pembangunan manusia yang tinggi akan meningkatkan produktivitas dan kreativitas sumber daya manusia. Dampak dari meningkatnya produktivitas dan kreativitas yang dapat diserap dan dikelola oleh sumber daya manusia, dimana sumberdaya ini sangat penting bagi pertumbuhan ekonomi.

2.11 Kemiskinan

Menurut BPS (2019) dalam mengukur kemiskinan suatu daerah mengunakan konsep kemampuan suatu individu dalam memenuhi kebutuhan dasar *(Basic need approach).* Pendekatan ini menganggap kemiskinan sebagai ketidakmampuan individu dalam memenuhi kebutuhan dasar makanan maupun non makanan berdasarkan pengeluaran yang dikeluarkan individu. Sehingga suatu individu dikatakan tergolong miskin apabila individu tersebut memiliki pengeluaran per-kapita dibawah garis kemiskinan.

Menurut Chamber yang dikutip oleh Suradi dalam (Mirza, 2012) menjabarkan bahwa kemiskinan sebagai suatu keadaan melarat dan ketidakberuntungan, suatu keadaan menurun *(deprivation)*, kemiskinan juga berkaitan dengan minimnya pendapatan dan harta serta lemah dalam fisik, isolasi, kerapuhan dan ketidakberdayaan suatu individu. Sedangkan menurut Amartya Sen yang dikutip oleh Suradi dalam (Mirza, 2012) menjelaskan bahwa kelaparan yang melanda suatu individu merupakan perspektif dari kemiskinan dengan ketidakmampuan dalam kehinaan dan ketidakmampuan dalam mendidik dan merawat kesehatan anak.

Hubungan kemiskinan dan Indeks Pembangunan Manusia dilihat dari indikator yang membangun Indeks Pembangunan Manusia seperti Pendidikan, Kesehatan dan Standar Hidup. Menurut Franciari & Sugiyanto (2012) menjelaskan bahwa indikator dari Indeks Pembangunan Manusia (IPM) seperti tingkat pendidikan dan tingkat kesehatan sangat erat hubungannya dengan kemiskinan.

Peningkatan dalam bidang kesehatan yang dilakukan pemerintah akan dapat meningkatkan kesehatan masyarakat dan anak-anak usia sekolah sehingga dapat menyerap pelajaran dengan baik. Tingkat pendidikan meningkatkan ketrampilan dan pengetahuan pada individu maka akan berdampak pada meningkatnya produktivitas dan pendapatan suatu individu. Hal ini berakibat pada meningkatnya pertumbuhan ekonomi yang kemudian akan menurunkan kemiskinan.

Kemiskinan sering dikaitkan dengan nilai upah yang didapat (pendapatan) oleh suatu individu, dimana pendapatan ini memiliki hubungan penting dengan Indeks Pembangunan Manusia. Pendapatan merupakan penentu utama dan hasil dari pembangunan manusia, dimana penduduk miskin akan mengunakan tenaganya dalam partisipasi dalam pertumbuhan ekonomi. Akan tetapi kemiskinan identik dengan kurangnya pendidikan, gizi dan kesehatan yang buruk sehingga menguranggi kapasitas dalam bekerja. Sehingga rendahnya Indeks Pembangunan Manusia dikarenakan penduduk miskin tidak dapat mengambil keutungan dari pendapatan produktif dari adanya pertumbuhan ekonomi suatu negara.

2.12 Rata Lama Sekolah

Menurut Trianggara dkk (2016) Rata- Rata lama sekolah merupakan salah satu indikator yang digunakan untuk menghitung komponen pendidikan. Pada Rata-Rata Lama Sekolah mempunyai bobot nilai sebesar satu pertiga dalam mengukur komponen pendidikan. Menurut Badan Pusat Statistik (2019) Rata lama sekolah adalah Jumlah tahun belajar penduduk usia 15 tahun ke atas yang telah diselesaikan dalam pendidikan formal. Indikator yang digunakan dalam penghitungan Rata lama sekolah antara lain: partisipasi sekolah, jenjang pendidikan, jenjang dan jenis pendidikan yang pernah /sedang diduduki, Ijasah tertinggi yang dimiliki dan Kelas tertinggi yang pernah/sedang diduduki.

Rata- rata lama sekolah digunakan untuk melihat kualitas penduduk dalam hal mengenyam pendidikan formal. Tinggi rendahnya nilai Rata- rata lama sekolah menunjukkan jenjang pendidikan yang pernah/sedang diduduki oleh seseorang. Semakin tinggi nilai Rata lama Sekolah maka semakin lama/tinggi jenjang pendidikan yang ditamatkannya dan Semakin rendah nilai Rata lama Sekolah maka semakin cepat/rendah jenjang pendidikan yang ditamatkannya.

2.13 Rata- Rata Lama Sakit

Menurut Faqihudin (2010) Rata- rata lama sakit merupakan salah satu indikator yang dapat mempengaruhi angka harapan hidup pada hasil Survey Sosial Ekonomi Nasional (Susenas) dalam bidang kesehatan. Menurut Badan Pusat Statistik (2019) Rata- rata lama sakit adalah rata –rata banyaknya hari sakit pada penduduk yang mengalami keluhan kesehatan. Rata- rata lama sakit yang dialami penduduk yakni selama 1 bulan terakhir.

Indikator Rata- rata lama sakit mengambarkan tingkat intensitas penyakit yang diderita penduduk. Indikator ini juga mengambarkan besarnya kerugian material karena penyakit yang diderita oleh penduduk. Rata-rata lama sakit digunakan untuk mengukur tingkat kesehatan masyarakat dan menunjukan seberapa serius keluhan yang diderita. Semakin besar nilai Rata-rata lama sakit maka semakin lama rata-rata lama hari sakit, semakin buruk tingkat kesehatan daerah dan semakin besar pula kerugian materiil yang dialami oleh penduduk.

2.14 Angka Partisipasi Sekolah

Menurut Badan Pusat Statistik (2019) angka partisipasi sekolah merupakan proporsi semua penduduk usia sekolah pada suatu kelompok umur tertentu terhadap penduduk dengan kelompok umur yang sesuai dan pendidikan non formal sejak tahun 2009 turut diperhitungkan. Angka partisipasi Sekolah yang semakin tinggi menunjukkan terbukanya kesempatan lebih besar dalam mengakses pendidikan dan semakin rendah angka partisipasi sekolah menunjukkan sempitnya kesempatan dalam mengakses pendidikan.

Menurut Astuti (2016) Angka Partisipasi Sekolah adalah salah satu ukuran daya serap lembaga pendidikan terhadap penduduk usia sekolah. Indikator angka partisipasi sekolah merupakan indikator dasar yang digunakan untuk melihat akses penduduk usia sekolah pada fasilitas pendidikan.