

Pemodelan Regresi Seemingly Unrelated Menggunakan Metode Maximum Likelihood pada Data Panel

Nurul Hikmah^{1*}, Raupong², Nasrah Sirajang³

¹²³Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Hadanuddin, Makassar, 60294, Indonesia

* Corresponding author, email: nrlhikmah22@gmail.com

Abstract

This study aims to model and predict the Human Development Index (HDI) values in South Sulawesi Province for the period 2014–2022 using a multiple linear regression approach with the Maximum Likelihood Estimation (MLE) method. Multiple linear regression analysis often encounters multicollinearity issues among independent variables; therefore, Principal Component Analysis (PCA) is employed as a dimensionality reduction technique to eliminate correlations among explanatory variables. In addition, due to the potential correlation of residuals among equations in a multivariate model, the Seemingly Unrelated Regression (SUR) approach is used, which is also estimated using the MLE method. The data utilized in this study is panel data, which offers advantages in obtaining more comprehensive and accurate information regarding the relationships between the analyzed variables. The estimation results of the SUR model indicate that variables such as Life Expectancy (UHH), Mean Years of Schooling (RLS), Expected Years of Schooling (HLS), and Adjusted Per Capita Expenditure have a significant influence on HDI across all districts/cities in South Sulawesi. One of the estimated equations from the SUR model is $y_{22t} = 81.44 + 0.670KU1_{22}$ which illustrates the relationship between the principal component and HDI in a specific region.

Keywords: Linear Regression, Maximum Likelihood Estimation, Seemingly Unrelated Regression, PCA, Human Development Index.

Abstrak

Penelitian ini bertujuan untuk memodelkan dan memprediksi nilai Indeks Pembangunan Manusia (IPM) di Provinsi Sulawesi Selatan pada periode 2014–2022 menggunakan pendekatan analisis regresi linier berganda dengan metode Maximum Likelihood Estimation (MLE). Analisis regresi linier berganda terkadang memiliki masalah multikolinearitas antar variabel independen, sehingga untuk mengatasinya digunakan metode *Principal Component Analysis* (PCA) sebagai teknik reduksi dimensi dan eliminasi korelasi antar variabel penjelas. Selain itu, karena kemungkinan terdapat korelasi residual antar persamaan dalam model multivariat, digunakan pendekatan Seemingly Unrelated Regression (SUR) yang juga diestimasi menggunakan metode MLE. Data yang digunakan dalam penelitian ini merupakan data panel, yang memberikan keunggulan dalam memperoleh informasi yang lebih komprehensif dan akurat terhadap hubungan antara variabel-variabel yang dianalisis. Hasil estimasi model SUR menunjukkan bahwa variabel Umur Harapan Hidup (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan Pengeluaran per Kapita Disesuaikan memiliki pengaruh signifikan terhadap IPM di seluruh Kabupaten/Kota di Sulawesi Selatan. Salah satu hasil estimasi model SUR adalah persamaan $y_{22t} = 81.44 + 0.670KU1_{22}$ yang menggambarkan hubungan antara komponen utama dan IPM di salah satu wilayah.

Kata Kunci: Regresi Linier, Maximum Likelihood Estimation, Seemingly Unrelated Regression, PCA, Indeks Pembangunan Manusia.

1. Pendahuluan

Analisis regresi adalah teknik statistika yang bertujuan untuk mengetahui pengaruh suatu variabel terhadap variabel lainnya. Variabel yang mempengaruhi disebut variabel independen (*independent variable*) dan variabel yang dipengaruhi disebut variabel dependen (*dependent variable*) [1]. Apabila pada regresi linier berganda terdapat variable-variabel independen yang saling berkorelasi, maka digunakan metode *Principal Component Analysis* (PCA) yang merupakan perwakilan dari variable-variabel independen. *Seemingly Unrelated Regresi* (SUR) adalah Teknik analisis regresi berguna ketika terjadi multikolinearitas antar kelompok-kelompok persamaannya [2]. Model SUR dapat diestimasi menggunakan metode *Maximum Likelihood Estimation* (MLE). MLE merupakan metode estimasi parameter model yang mencari nilai parameter yang memaksimalkan fungsi *likelihood*. Dalam konteks SUR, MLE digunakan untuk mengevaluasi parameter yang terkait dengan masing-masing persamaan regresi yang saling terkait dalam konteks model SUR [3].

Data yang digunakan yaitu data panel. Salah satu keuntungan menggunakan data panel yaitu dapat memperoleh informasi yang lebih detail dan lengkap mengenai hubungan antara variabel independen dan variabel dependen dari data yang dikumpulkan dari waktu ke waktu [4].

Penelitian sebelumnya dilakukan Pusakasari (2015), mengenai model data panel yang dipilih yaitu *Fixed Effect Model* (FEM) dan digunakan pembobot *cross-section* SUR, karena pengamatan *Gross Domestic Product* (GDP) mengandung heteroskedastisitas dan korelasi antar unit *cross-section*. Selanjutnya penelitian oleh Hufaini dkk. (2020), mengenai model regresi yang diperoleh dari estimasi *maximum likelihood* dengan pendekatan metode *within group* pada data panel. Hasil penelitiannya yaitu estimasi *maximum likelihood* mampu menjelaskan efek perbedaan wilayah serta mengetahui perbedaan intersep untuk efek tahun.

2. Material dan Metode

2.1 Regresi Data Panel

Analisis regresi data panel adalah analisis regresi yang didasarkan pada data panel untuk mengamati hubungan antara satu variabel dependen dengan satu atau lebih variabel independen. Bentuk umum data panel ditunjukkan pada Tabel 1 [5].

Tabel 1. Bentuk Umum Data Panel

Individu (<i>i</i>)	Periode (<i>t</i>)	Y_{it}	X_{1it}	X_{2it}	...	X_{pit}
1	1	Y_{11}	X_{111}	X_{211}	...	X_{p11}
	2	Y_{12}	X_{112}	X_{212}	...	X_{p12}

	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	T	Y_{1T}	X_{11T}	X_{21T}	$\dots$	X_{p1T}
2	1	Y_{21}	X_{121}	X_{221}	$\dots$	X_{p21}
	2	Y_{22}	X_{122}	X_{222}	$\dots$	X_{p22}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	T	Y_{2T}	X_{11T}	X_{21T}	$\dots$	X_{p1T}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	1	Y_{N1}	X_{1N1}	X_{2N1}	$\dots$	X_{pN1}
	2	Y_{N2}	X_{1N2}	X_{2N2}	$\dots$	X_{pN2}
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	T	Y_{NT}	X_{1NT}	X_{2NT}	$\dots$	X_{pNT}

Model umum regresi data panel dapat dinyatakan dalam persamaan berikut [6]:

$$y_{it} = \beta_{0i} + \sum_{k=1}^p \beta_k x_{kit} + \varepsilon_{it}; \quad (1)$$

$$i = 1, 2, \dots, N; t = 1, 2, \dots, T; k = 1, 2, \dots, p$$

dengan:

N : jumlah individu (*cross section*)

T : jumlah periode waktu (*time series*)

y_{it} : variabel dependen pada individu ke- i untuk periode ke- t

β_{0i} : koefisien intersep (konstanta) ke- i

β_{ik} : koefisien regresi variabel independen ke- k ,

x_{it} : variabel independent ke- k pada individu ke- i untuk periode ke- t

ε_{it} : residual individu ke- i dan waktu ke- t

2.2 Uji Asumsi Klasik

a. Uji Normalitas

Salah satu metode yang umum digunakan untuk menguji asumsi normalitas residual adalah uji Shapiro-Wilk dengan hipotesis uji sebagai berikut [7]:

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Statistik Uji:

$$W = \frac{1}{S^2} \left[\sum_{i=1}^n \alpha_i (e_{n-i+1} - e_i) \right]^2$$

$$S^2 = \sum_{i=1}^n (e_i - \bar{e})^2 \quad (2)$$

dengan:

α_i : Koefisien *Shapiro-Wilk test*

e_{n-i+1} : Residual ke $n - i + 1$
 e_i : Residual ke- i
 $\bar{e}$: Rata-rata residual
 S^2 : Varians residual

Terima H_0 jika statistik uji $W > W(n, \alpha)$ maka residual berdistribusi normal.

b. Uji Multikolinearitas

Uji multikolinearitas bertujuan menguji apakah model regresi terdapat korelasi antar variabel independen atau tidak. Model regresi yang baik seharusnya tidak terjadi korelasi diantara variabel independen. Multikolinearitas dapat dilihat dengan menggunakan rumus sebagai berikut:

$$VIF = \frac{1}{1 - R_k^2} \quad (3)$$

Dengan R_k merupakan koefisien determinasi dari model. Dikatakan terjadi multikolinearitas apabila nilai $VIF > 10$. Jika terjadi multikolinearitas maka data saling tumpah tindih sehingga dapat menyebabkan informasi kurang akurat [8].

2.3 Principal Component Analysis

Tujuan PCA adalah untuk mengekstrak informasi penting dari data dan mengaturnya ke variabel baru yang disebut principal component (komponen utama). Ada tiga kriteria dalam pemilihan komponen utama yang digunakan yaitu dengan melihat nilai eigen yang lebih besar dari satu, melihat sudut pada scree plot yang menunjukkan perubahan nilai eigen yang besar, dan proporsi variansi kumulatif mencapai 80% [9].

Nilai $Z = (Z_1, Z_2, \dots, Z_N)$ adalah nilai $X = (X_1, X_2, \dots, X_N)$ yang telah terstandarisasi menjadi variabel baru pada $KU = (KU_1, KU_2, \dots, KU_g)$.

$$Z_i = \frac{X_i - \mu_i}{\sigma_i} \quad (4)$$

dengan:

X_i : nilai variabel ke- i pada objek, $i = 1, 2, \dots, N$

μ_i : nilai rata-rata variabel ke- i

σ_i : nilai standar deviasi variabel ke- i

Z_i : nilai standar i ,

maka diperoleh:

$$U_f = b_{1f}Z_1 + b_{2f}Z_2 + \dots + b_{Nf}Z_N \quad (5)$$

dengan demikian, persamaan PCA dapat ditulis menjadi ,

$$\begin{aligned} KU_1 &= b_{11}Z_1 + b_{21}Z_2 + \dots + b_{N1}Z_N \\ KU_2 &= b_{12}Z_1 + b_{22}Z_2 + \dots + b_{N2}Z_N \\ &\vdots \\ KU_g &= b_{1g}Z_1 + b_{2g}Z_2 + \dots + b_{Ng}Z_N \end{aligned}$$

dengan:

KU_f : komponen utama ke- f , $f = 1, 2, \dots, g$

b_{Nf} : nilai vektor eigen variabel N pada KU_f

2.4 Seemingly Unrelated Regression

Model Seemingly Unrelated Regression (SUR) merupakan model regresi yang pertama kali diperkenalkan oleh Zellner pada tahun 1996. SUR adalah suatu model yang terdiri dari beberapa persamaan yang memiliki kaitan satu sama lain karena residual antar persamaan yang berbeda saling berkorelasi. Model SUR dapat dituliskan dalam persamaan model regresi linier [11] sebagai berikut:

$$\begin{aligned} y_{1t} &= \beta_{01} + \beta_{11}x_{11t} + \beta_{12}x_{21t} + \dots + \beta_{1k}x_{k1t} + \varepsilon_{1t} \\ y_{2t} &= \beta_{02} + \beta_{21}x_{12t} + \beta_{22}x_{22t} + \dots + \beta_{2k}x_{k2t} + \varepsilon_{2t} \\ &\vdots \\ y_{Nt} &= \beta_{0N} + \beta_{N1}x_{1Nt} + \beta_{N2}x_{2Nt} + \dots + \beta_{Nk}x_{kNt} + \varepsilon_{Nt} \end{aligned} \quad (6)$$

Y_{it} adalah variabel tak independen untuk persamaan regresi ke- i pada pengamatan ke- t , X_{kit} adalah variabel independen untuk persamaan regresi ke- i pada pengamatan ke- t , β_{ik} adalah parameter regresi untuk persamaan regresi ke- i dan ε_{it} adalah komponen acak galat untuk persamaan ke- i . Secara khusus, sistem persamaan regresi SUR dapat ditulis seperti berikut:

$$Y_i = X_i\beta_i + \varepsilon_i ; \quad (i = 1, 2, \dots, N) \quad (7)$$

dapat dituliskan sebagai lambang matriks sebagai berikut

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{pmatrix} = \begin{pmatrix} X_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & X_2 & \ddots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & X_N \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{pmatrix}$$

dimana $Y_i = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{pmatrix}$, $X_i = \begin{pmatrix} X_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & X_2 & \ddots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & X_N \end{pmatrix}$,

$$\beta_i = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{pmatrix}, \quad \varepsilon_i = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{pmatrix}$$

dengan $Y_1 = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1t} \end{bmatrix}$ $Y_2 = \begin{bmatrix} y_{21} \\ y_{22} \\ \vdots \\ y_{2t} \end{bmatrix}$... $Y_N = \begin{bmatrix} y_{N1} \\ y_{N2} \\ \vdots \\ y_{Nt} \end{bmatrix}$

$$X_1 = \begin{bmatrix} 1 & x_{111} & x_{211} & \dots & x_{k11} \\ 1 & x_{112} & x_{212} & \dots & x_{k12} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{1NT} & x_{2NT} & \dots & x_{kNT} \end{bmatrix} \quad X_2 = \begin{bmatrix} 1 & x_{121} & x_{221} & \dots & x_{k21} \\ 1 & x_{122} & x_{222} & \dots & x_{k22} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{1NT} & x_{2NT} & \dots & x_{kNT} \end{bmatrix} \dots$$

$$X_N = \begin{bmatrix} 1 & x_{1N1} & x_{2N1} & \dots & x_{kN1} \\ 1 & x_{1N2} & x_{2N2} & \dots & x_{kN2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{1NT} & x_{2NT} & \dots & x_{kNT} \end{bmatrix}$$

$$\beta_1 = \begin{bmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1t} \end{bmatrix} \quad \beta_2 = \begin{bmatrix} \beta_{20} \\ \beta_{21} \\ \beta_{22} \\ \vdots \\ \beta_{2t} \end{bmatrix} \quad \dots \quad \beta_N = \begin{bmatrix} \beta_{N0} \\ \beta_{N1} \\ \beta_{N2} \\ \vdots \\ \beta_{Nt} \end{bmatrix}$$

$$\varepsilon_1 = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \vdots \\ \varepsilon_{1t} \end{bmatrix} \quad \varepsilon_2 = \begin{bmatrix} \varepsilon_{21} \\ \varepsilon_{22} \\ \vdots \\ \varepsilon_{2t} \end{bmatrix} \quad \dots \quad \varepsilon_N = \begin{bmatrix} \varepsilon_{N1} \\ \varepsilon_{N2} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}$$

Persamaan Y_i adalah vektor pengamatan berukuran $NT \times 1$; X_i adalah vektor variabel independen berukuran $(NT \times (p + 1))$; β_i adalah vektor parameter regresi ke- i pada model SUR berukuran $((p + 1) \times 1)$; dan ε_i adalah vektor kolom galat berukuran $(NT \times 1)$.

2.5 Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) adalah salah satu metode estimasi yang paling sering digunakan dalam analisis regresi linier berganda. Metode ini memperkirakan parameter dengan memaksimalkan fungsi likelihood dari model regresi, sehingga dapat menghasilkan estimasi parameter yang akurat dan efisien. Fungsi *likelihood* dari banyaknya variabel acak $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ didefinisikan sebagai fungsi kepadatan bersama dari banyaknya variabel acak fungsi kepadatan bersama $f(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n; \beta)$, yang mempertimbangkan fungsi dari β jika $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ adalah $f(\varepsilon_1; \beta)f(\varepsilon_2; \beta) \dots f(\varepsilon_n; \beta)$. Estimasi parameter β metode MLE dengan asumsi nilai ε berdistribusi $\varepsilon \sim N(\mu, \sigma^2)$, maka fungsi kepadatan bersama sebagai berikut [11]:

$$f(\boldsymbol{\varepsilon}, \boldsymbol{\beta}) = \prod_{i=1}^n f(\varepsilon_i) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n \varepsilon_i^2\right) \quad (8)$$

Sehingga fungsi likelihood dapat ditulis sebagai berikut:

$$L(\boldsymbol{\beta}; \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n \varepsilon_i^2\right)$$

karena $\sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$, maka fungsi likelihood dapat ditulis sebagai berikut:

$$L(\boldsymbol{\beta}, \boldsymbol{\varepsilon}) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma^2} (\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon})\right) \quad (9)$$

Untuk memperoleh penduga parameter $\hat{\boldsymbol{\beta}}$, fungsi likelihood dilogaritmakan dan diturunkan terhadap parameter-parameternya lalu disamakan dengan nol.

$$l = \ln L(\boldsymbol{\beta}, \boldsymbol{\varepsilon})$$

$$l = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon}$$

$$l = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

$$l = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} (\mathbf{Y}^T \mathbf{Y} - 2\boldsymbol{\beta}^T \mathbf{X}^T \mathbf{Y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X} \boldsymbol{\beta})$$

$$(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X} \boldsymbol{\beta} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{Y} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X} \boldsymbol{\beta}$$

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

$$\left. \frac{dl}{d\boldsymbol{\beta}} \right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} = 0$$

$$\mathbf{X}^T \mathbf{Y} - \mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = 0$$

$$\mathbf{X}^T \mathbf{X} \hat{\boldsymbol{\beta}} = \mathbf{X}^T \mathbf{Y} \text{ dikali } (\mathbf{X}^T \mathbf{X})^{-1}$$

Maka:

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

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

diperoleh persamaan berikut:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (10)$$

Sehingga, diperoleh penduga parameter model SUR dengan metode MLE adalah sebagai berikut:

$$\hat{\boldsymbol{\beta}}_i = (\mathbf{X}_i^T \mathbf{X}_i)^{-1} \mathbf{X}_i^T \mathbf{Y}_i \quad (11)$$

2.6 Uji Signifikansi Parameter

a. Uji Simultan

Uji simultan (uji F) dilakukan untuk menguji hasil estimasi model regresi apakah variabel independen secara bersama-sama memiliki pengaruh terhadap variabel dependen. Dengan hipotesis sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0, k = 1, 2, \dots, p$$

Statistik uji:

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (12)$$

H_0 ditolak jika $F_{hitung} > F_{tabel}(F_{\alpha; (p-1, n-p)})$, sehingga hubungan antar variabel dependen dan independen berpengaruh signifikan.

b. Uji Parsial

Uji parsial (uji t) digunakan untuk melihat pengaruh secara individu variabel independen dalam model. Hipotesis sebagai berikut [13].

$$H_0 : \beta_k = 0 \text{ (variabel independen tidak berpengaruh signifikan terhadap variabel dependen)}$$

$$H_1 : \beta_k \neq 0 \text{ (variabel independen berpengaruh signifikan terhadap variabel dependen)}$$

Statistik uji:

$$t = \frac{b_1}{s(b_1)} \quad (13)$$

H_0 ditolak jika $|t_{hitung}| > t_{\frac{\alpha}{2}, n-p-1}$ atau $p\text{-value} < \alpha$ yang artinya variabel independen berpengaruh signifikan terhadap variabel dependen

3. Hasil dan Diskusi

3.1 Analisis Deskriptif

Data pada penelitian ini dideskripsikan berdasarkan nilai minimum, nilai maksimum, dan nilai rata-rata. Hasil analisis deskriptif yang dilakukan disajikan pada Tabel 2 berikut:

Tabel 2. Statistik Deskriptif

Variabel	Jumlah Data	Minimum	Maksimum	Rata-rata
Y	216	61.45	83.12	69.48
X_1	216	65.37	73.72	69.08
X_2	216	5.63	11.55	7.89
X_3	216	11.31	15.59	13.08
X_4	216	6214	17406	10533.77

Tabel 2 menunjukkan bahwa data yang digunakan sebanyak 216 observasi yang diperoleh dari 24 Kabupaten/Kota di Sulawesi Selatan dengan rentang waktu 9 tahun yaitu dari tahun 2014 hingga tahun 2022. Rata-rata nilai IPM di Sulawesi Selatan dari tahun 2014-2022 sebesar 69.48. Nilai IPM terendah sebesar 61.45 berada pada Kabupaten Jeneponto pada tahun 2014, sedangkan nilai IPM tertinggi sebesar 83.12 berada pada Kota Makassar pada tahun 2022.

Umur harapan hidup di setiap Kabupaten/Kota di Provinsi Sulawesi Selatan diduga berpengaruh dalam menentukan nilai IPM. Berdasarkan Tabel 2, diketahui bahwa rata-rata umur harapan hidup di Sulawesi Selatan adalah 69.08 yang berarti bahwa setiap satu anak lahir di Sulawesi Selatan diharapkan dapat bertahan hidup selama 69.08 tahun. Nilai AHH terendah sebesar 65.37 berada pada Kabupaten Pangkajene dan Kepulauan pada tahun 2014, sedangkan nilai AHH tertinggi sebesar 73.72 berada pada Kabupaten Tana Toraja pada tahun 2022.

Rata-rata lama sekolah di setiap Kabupaten/Kota di Provinsi Sulawesi Selatan diduga berpengaruh dalam menentukan IPM. Berdasarkan Tabel 2, diketahui bahwa rata-rata RLS di Sulawesi Selatan adalah 7.89 tahun. Nilai RLS terendah 5.63 berada pada Kabupaten Jeneponto pada tahun 2014, sedangkan nilai RLS tertinggi sebesar 11.55 berada pada Kota Makassar pada tahun 2022.

Harapan lama sekolah di setiap Kabupaten/Kota di Provinsi Sulawesi Selatan diduga berpengaruh dalam menentukan nilai IPM. Berdasarkan Tabel 2, diketahui bahwa rata-rata harapan lama sekolah di Sulawesi Selatan adalah 13.08 tahun. Nilai harapan lama sekolah terendah sebesar 11.31 berada pada Kabupaten Takalar pada tahun 2014, sedangkan nilai harapan lama sekolah tertinggi sebesar 15.59 berada pada Kota Makassar pada tahun 2022.

Pengeluaran perkapita di setiap Kabupaten/Kota di Provinsi Sulawesi Selatan diduga berpengaruh dalam menentukan nilai IPM. Berdasarkan Tabel 2, diketahui bahwa rata-rata pengeluaran perkapita di Sulawesi Selatan adalah 10533.77 juta rupiah. Nilai pengeluaran perkapita terendah sebesar 6.214 juta rupiah berada pada Kabupaten Tana Toraja pada tahun 2014, sedangkan nilai pengeluaran perkapita tertinggi sebesar 17406 juta rupiah berada pada Kota Makassar pada tahun 2022.

3.2 Analisis Deskriptif

a. Uji Normalitas

Pada penelitian ini, uji normalitas dilakukan untuk menguji apakah residual memenuhi asumsi normalitas untuk mengevaluasi kebaikan model regresi. Metode yang digunakan yaitu uji *Shapiro-Wilk*. Adapun hasil dari uji normalitas sebagai berikut.

Hipotesis:

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Hasil yang diperoleh dari uji *Shapiro-Wilk* sebagai berikut.

Tabel 3. Test of Normality

	<i>Shapiro-Wilk</i>		
	Statistik	df	Sig.
y	0.950	9	0.695

Tabel 3, nilai uji *Shapiro-Wilk* yang telah diperoleh yaitu $\text{Sig} = 0.695 > 0.05$, maka dapat disimpulkan bahwa H_0 diterima dengan artian data berdistribusi normal.

b. Uji Multikolinearitas

Pada penelitian ini, uji multikolineritas dilakukan untuk menguji apakah model regresi terdapat korelasi antar variabel independen atau tidak. Adapun hasil dari uji multikolineritas sebagai berikut:

Tabel 4. Nilai Uji Multikolinearitas

Kabupaten/Kota	VIF			
	X_1	X_2	X_3	X_4
Kepulauan Selayar	73.019	25.810	4.604	23.910
Bulukumba	28.810	52.890	27.164	25.098
Bantaeng	89.661	9.560	4.856	104.453
Jeneponto	39.460	35.059	23.133	118.445
Takalar	60.684	48.223	9.215	36.387
Gowa	17.948	48.125	25.807	35.020
Sinjai	36.695	15.507	5.985	33.334
Maros	17.304	6.677	4.666	18.622
Pangkep	173.610	12.017	19.694	79.035
Barru	102.418	20.667	9.131	63.759
Bone	28.629	16.204	33.100	21.136
Soppeng	31.484	18.286	36.361	140.102
Wajo	25.951	7.386	5.447	15.372
Sidrap	35.189	18.151	21.268	19.350
Pinrang	177.390	18.926	274.042	64.696
Enrekang	8.107	12.767	10.121	31.881
Luwu	171.553	111.972	4.231	38.631
Tana Toraja	42.567	7.588	19.815	15.152
Luwu Utara	39.440	22.465	7.108	21.266

Luwu Timur	50.418	19.185	2.885	46.335
Toraja Utara	41.493	6.204	4.392	26.497
Makassar	6.499	20.358	10.712	39.853
Parepare	93.023	14.441	1.917	91.872
Palopo	27.701	59.041	42.187	99.532

Data dikatakan multikolineritas apabila nilai dari $VIF > 10$. Oleh karena itu, pada Tabel 4 yang diperoleh disimpulkan bahwa terdapat multikolineritas pada setiap Kabupaten/Kota di Sulawesi Selatan karena memiliki nilai $VIF > 10$. Adanya multikolineritas menyebabkan informasi yang disajikan menjadi tidak akurat. Solusi dalam permasalahan tersebut dapat diatasi dengan metode PCA.

3.3 *Principal Component Analysis*

Berdasarkan uji asumsi multikolineritas yang telah dibahas sebelumnya, diperoleh bahwa data yang digunakan mengandung multikolineritas. Oleh sebab itu untuk mengatasi multikolineritas tersebut, maka dilakukan analisis komponen utama. Adapun hasil yang diperoleh menunjukkan bahwa nilai $\lambda > 1$ atau nilai variansi kumulatif mencapai 80% sebagai berikut:

Tabel 5. Proporsi Keragaman Kumulatif untuk Setiap Kabupaten/Kota

Kab/Kota	Komponen Utama	λ	Cumulative
Kepulauan Selayar	KU_1	3.699	0.925
Bulukumba	KU_2	3.915	0.979
Bantaeng	KU_3	3.731	0.933
Jeneponto	KU_4	3.884	0.971
Takalar	KU_5	3.769	0.942
Gowa	KU_6	3.873	0.968
Sinjai	KU_7	3.701	0.925
Maros	KU_8	3.622	0.906
Pangkep	KU_9	3.750	0.937
Barru	KU_{10}	3.825	0.956
Bone	KU_{11}	3.864	0.966
Soppeng	KU_{12}	3.875	0.969
Wajo	KU_{13}	3.648	0.912
Sidrap	KU_{14}	3.805	0.951
Pinrang	KU_{15}	3.899	0.975
Enrekang	KU_{16}	3.650	0.913
Luwu	KU_{17}	3.700	0.925
Tana Toraja	KU_{18}	3.761	0.940

Luwu Utara	KU_{19}	3.816	0.954
Luwu Timur	KU_{20}	3.564	0.891
Toraja Utara	KU_{21}	3.409	0.852
Makassar	KU_{22}	3.682	0.921
Parepare	KU_{23}	3.446	0.861
Palopo	KU_{24}	3.938	0.985

3.4 Penduga Model SUR Menggunakan MLE

Pada penelitian ini, penduga model SUR diduga dengan menerapkan metode MLE. Data yang digunakan yaitu hasil dari pembentukan nilai komponen utama yang terbentuk. Berikut hasil dari nilai penduga parameter MLE.

Tabel 6. Nilai Penduga Parameter MLE

Kabupaten/Kota	$\hat{\beta}_0$	$\hat{\beta}_1$
Kepulauan Selayar	66.084	0.842
Bulukumba	67.699	0.895
Bantaeng	67.700	0.695
Jeneponto	63.202	0.698
Takalar	66.043	0.855
Gowa	68.774	0.839
Sinjai	66.271	0.801
Maros	68.852	0.780
Pangkep	67.849	0.637
Barru	69.947	0.627
Bone	64.822	0.830
Soppeng	67.323	0.882
Wajo	68.416	0.654
Sidrap	70.314	0.665
Pinrang	70.433	0.555
Enrekang	71.722	0.731
Luwu	69.543	0.700
Tana Toraja	67.548	0.864
Luwu Utara	68.761	0.628
Luwu Timur	72.003	0.762
Toraja Utara	68.384	0.765
Makassar	81.440	0.670
Parepare	77.173	0.516
Palopo	77.301	0.553

Setelah diperoleh nilai $\hat{\beta}$ menggunakan metode MLE, maka diperoleh estimasi model SUR menggunakan MLE, dalam hal ini yang ditampilkan yaitu untuk Kota Makassar dengan persentase tertinggi dari 24 Kabupaten/Kota di Provinsi Sulawesi Selatan.

$$y_{22t} = 81.44 + 0.670KU1_{22}$$

Persamaan diatas menunjukkan bahwa nilai Indeks Pembangunan Manusia (IPM) akan bertambah sebesar 0.670 jika variabel pada komponen utama $KU1$ bertambah satu satuan. Artinya, nilai $KU1$ telah mewakili variable-variabel independen yaitu umur harapan hidup (X_1), rata-rata lama sekolah (X_2), harapan lama sekolah (X_3), dan pengeluaran perkapita (X_4).

3.5 Uji Signifikansi Parameter

Pada penelitian ini, yang digunakan pada uji signifikansi parameter yaitu uji parsial untuk melihat pengaruh variabel independen dalam model. Adapun hasil dari penelitian ini adalah sebagai berikut:

Hipotesis:

$$H_0: \beta_{i1} = 0$$

$$H_1: \beta_{i1} \neq 0, i = 1, 2, \dots, 24$$

Hasil untuk Kota Makassar nilai $t_{hitung} = 3041.2645$ lebih besar dari $t_{tabel}(t_{0.025;7}) = 2.841$. Oleh karena itu, dapat disimpulkan bahwa H_0 ditolak berarti ada pengaruh skor komponen utama pertama terhadap IPM yang artinya 92.1% variabel-variabel independen umur harapan hidup (X_1), rata-rata lama sekolah (X_2), harapan lama sekolah (X_3), dan pengeluaran perkapita (X_4) berpengaruh signifikan terhadap variabel dependen indeks pembangunan manusia (Y).

4. Kesimpulan

Berdasarkan analisis dan pembahasan pada bab sebelumnya, dapat ditarik kesimpulan sebagai berikut:

1. Model *Seemingly Unrelated Regression* (SUR) menggunakan metode *Maximum Likelihood Estimation* (MLE) untuk semua Kabupaten/Kota di Sulawesi Selatan sebagai berikut:

$$y_{1t} = 66.008 + 0.842KU1_1$$

$$y_{2t} = 67.699 + 0.895KU1_2$$

$$y_{3t} = 67.7 + 0.695KU1_3$$

$$y_{4t} = 63.202 + 0.698KU1_4$$

$$y_{5t} = 66.043 + 0.855KU1_5$$

$$y_{6t} = 68.774 + 0.840KU1_6$$

$$y_{7t} = 66.271 + 0.801KU1_7$$

$$y_{8t} = 68.852 + 0.780KU1_8$$

$$y_{9t} = 67.845 + 0.637KU1_9$$

$$y_{10t} = 69.947 + 0.627KU1_{10}$$

$$y_{11t} = 64.822 + 0.830KU1_{11}$$

$$y_{12t} = 67.323 + 0.882KU1_{12}$$

$$y_{13t} = 68.416 + 0.654KU_{13}$$

$$y_{14t} = 70.314 + 0.665KU_{14}$$

$$y_{15t} = 70.433 + 0.555KU_{15}$$

$$y_{16t} = 71.722 + 0.731KU_{16}$$

$$y_{17t} = 69.543 + 0.700KU_{17}$$

$$y_{18t} = 67.548 + 0.864KU_{18}$$

$$y_{19t} = 68.761 + 0.628KU_{19}$$

$$y_{20t} = 72.003 + 0.762KU_{20}$$

$$y_{21t} = 63.384 + 0.765KU_{21}$$

$$y_{22t} = 81.44 + 0.670KU_{22}$$

$$y_{23t} = 77.172 + 0.516KU_{23}$$

$$y_{24t} = 77.301 + 0.553KU_{24}$$

$$t = 1, 2, 3, 4, 5, 6, 7, 8, 9$$

2. Faktor-faktor umur harapan hidup (X_1), rata-rata lama sekolah (X_2), harapan lama sekolah (X_3), dan pengeluaran per kapita disesuaikan (X_4) berpengaruh terhadap indeks pembangunan manusia (Y) di semua Kabupaten/Kota Provinsi Sulawesi Selatan.

Daftar Pustaka

- [1] Sari, A. Q., Sukestiyarno, Y. L., & Agoestanto, A. Batasan Prasyarat Uji Normalitas dan Uji Homogenitas pada Model Regresi Linear. *Unnes Journal of Mathematics*, 6(2), 168–177, 2017.
- [2] Zellner, A. An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, 57(298), 348–368, 1962.
- [3] Wooldridge, J. M. *Introductory Econometrics. A Modern Approach (Fifth Edition)*. South-Western, 2012.
- [4] Ratnasari, N., Kencana, I., & Gandhiadi, G. K. Aplikasi Regresi Data Panel dengan Pendekatan Fixed Effect Model (studi kasus: PT PLN Gianyar). *E-jurnal matematika*, 3(1), 1–7, 2014.
- [5] Ghozi, S., & Hermansyah, H. Analisis Regresi Data Panel Profitabilitas Bank Pembangunan Daerah (BPD) di Indonesia. *Jurnal Matematika*, 8(1), 1–12, 2018.
- [6] Baltagi, B. H. (2005). *Econometric Analysis of Panel Data (Third edition)*. John Wiley & Sons, Ltd, 2005.
- [7] Syamsudin, R., & Wachidah, L. Pengujian Asumsi Homokedastisitas Regresi Linear Berganda Menggunakan RCEV Test Studentized Residual pada Data Pendapatan Asli Daerah (PAD) Kabupaten/Kota Jawa Barat Tahun 2018. *Prosiding Statistika*. 6(2). 9-16, 2020.
- [8] Gujarati, D. N., & Porter, D. C. *Basic Econometrics (Fifth Edition)*. McGraw-Hill, 2009.

- [9] Sari, N., Yasin, H., & Prahutama, A. Geographically Weighted Regression Principal Component Analysis (GWRPCA) Pada Pemodelan Pendapatan Asli Daerah Di Jawa Tengah. *Jurnal Gaussian*, 5(4)717-726, 2016.
- [10] Abdi, H., & Williams, L. J. Principal Component Analysis. Wiley Interdisciplinary Reviews: *Computational Statistics*, 2(4), 433–459, 2010.
- [11] Moon, H. R., & Perron, B. Seemingly Unrelated Regressions. *The new palgrave dictionary of economics*, 1(9), 2006.
- [12] Hidayat, M. J., Hadi, A. F., & Anggraeni, D. Analisis Regresi Data panel Terhadap Indeks Pembangunan Manusia (IPM) Jawa Timur 2006-2015 (Panel Data Regression Analysis of the East Java Human Development Index (HDI) 2006-2015). *Dalam Majalah Ilmiah Matematika dan Statistika*, 18, 2018.
- [13] Fitria, S. F. Analisis Regresi Data Panel Pengaruh PDRB, Indeks Pembangunan Manusia, dan Tingkat Pengangguran Terbuka terhadap Jumlah Kemiskinan di Kabupaten/Kota di Jawa Barat pada Tahun 2013-2020. *Jurnal Riset Matematika*, 1(2), 119–128, 2021.