

# Analisis Regresi Komponen Utama Untuk Mengatasi Multikolinieritas Pada Faktor-Faktor Yang Mempengaruhi Indeks Pembangunan Manusia

*(Principal Component Regression to Overcome Multicollinearity in Factors That Influence the Human Development Index)*

Azka Fariz Hidayatullah<sup>[1]</sup>, Dede Saputra<sup>[2]</sup>, Filzah Inarah<sup>[3]</sup>, Isma Evita<sup>[4]</sup>, Muhammad Fadillah<sup>[5]</sup>, Lisa Harsyiah<sup>[6]\*</sup>

<sup>[1], [2], [3], [4], [5]</sup>Program Studi Matematika, Universitas Mataram

E-mail: [dedesaputr2@gmail.com](mailto:dedesaputr2@gmail.com)

<sup>[6]</sup>Program Studi Statistika, Universitas Mataram

E-mail: [lisa\\_harsyiah@unram.ac.id](mailto:lisa_harsyiah@unram.ac.id)

## KEYWORDS:

Human Development Index, Multicollinearity, Principal Component Regression

## KATA KUNCI:

DES Brown, DES Holt, Golden Section, Peramalan

## ABSTRACT

West Nusa Tenggara and East Nusa Tenggara provinces are among the provinces with low human development indexes. There are seven factors used in this study that are considered to affect the human development index in the two provinces, namely gross regional domestic product, poor population, open unemployment rate, population, life expectancy, labor force and average years of schooling. The method used by researchers in overcoming multicollinearity in this study is principal component regression. Therefore, this study aims to apply principal component regression in overcoming the problem of multicollinearity on the the effect of gross regional domestic product, poor population, open unemployment rate, population, life expectancy, labor force and average years of schooling on the human development index. Based on the results of the analysis that has been carried out, the principle component regression model is obtained as follows  $Y = 42.548 + 0.00000991X_1 + 0.0371X_2 + 0.343X_3 + 0.000005949X_4 + 0.2532 + 0.000012947X_6 + 0.1348X_7$  . With the coefficient of determination ( $R^2$ ) 0.834.

## ABSTRAK

Provinsi Nusa Tenggara Barat dan Nusa Tenggara Timur masuk dalam provinsi yang memiliki indeks pembangunan manusia yang rendah. Terdapat tujuh faktor yang digunakan dalam penelitian ini yang dianggap mempengaruhi indeks pembangunan manusia di dua provinsi tersebut yaitu pembangunan manusia yaitu produk domestik regional bruto, penduduk miskin, tingkat pengangguran terbuka, jumlah penduduk, angka harapan hidup, jumlah angkatan kerja dan rata-rata lama sekolah. Metode yang digunakan peneliti dalam mengatasi multikolinieritas adalah regresi komponen utama. Oleh karena itu penelitian ini bertujuan untuk menerapkan regresi komponen utama dalam mengatasi masalah multikolinieritas pada pengaruh produk domestik regional bruto, penduduk miskin, tingkat pengangguran terbuka, jumlah penduduk, angka harapan hidup, jumlah angkatan kerja dan rata-rata lama sekolah terhadap indeks pembangunan manusia. Berdasarkan hasil analisis yang telah dilakukan diperoleh model regresi komponen utama yaitu  $Y = 42.548 + 0.00000991X_1 + 0.0371X_2 + 0.343X_3 + 0.000005949 X_4 + 0.2532 + 0.000012947X_6 + 0.1348X_7$  dengan nilai koefisien determinasi ( $R^2$ ) sebesar 0.834.

## 1. PENDAHULUAN

Pembangunan manusia merupakan suatu indikator yang merepresentasikan keberhasilan dalam membangun hidupnya yang dianggap berharga yang dapat merasakan

kehidupan yang panjang dan sehat, berpengetahuan dan mempunyai akses mengenai sumber-sumber yang dibutuhkan dalam menjalani kehidupan yang layak (Pratowo, 2012). Pembangunan manusia merupakan

salah satu indikator yang dianggap krusial yang mencerminkan keberhasilan dalam pembangunan ekonomi di suatu daerah. Oleh karena itu, dalam perencanaan pembangunan ekonomi saat ini pembangunan manusia menjadi prioritas yang terus ditekankan (Si'lang, Hasid & Priyagus, 2019). Indeks pembangunan manusia adalah suatu indikator yang digunakan untuk mengukur tingkat keberhasilan suatu daerah dalam proses pembangunan kualitas yang berfokus pada manusia (Rosyadah, 2021).

Indonesia saat ini memiliki fokus dalam proses pembangunan dan peningkatan terhadap pertumbuhan ekonomi. Dalam memenuhi tujuan tersebut perlu ditinjau indeks pembangunan manusia di setiap daerah di Indonesia. Berdasarkan data dari Badan Perencanaan Pembangunan Nasional dan Badan Pusat Statistik Provinsi Nusa Tenggara Barat dan Nusa Tenggara Timur masuk dalam provinsi yang memiliki indeks pembangunan manusia rendah dengan skor berturut-turut yaitu 68,25 dan 65,19. Oleh karena itu dalam meningkatkan indeks pembangunan manusia di Provinsi tersebut perlu dikaji lebih lanjut faktor-faktor yang mempengaruhinya. Penelitian ini menggunakan tujuh faktor yang dianggap mempengaruhi indeks pembangunan manusia yaitu produk domestik regional bruto, penduduk miskin, tingkat pengangguran terbuka, jumlah penduduk, angka harapan hidup, jumlah angkatan kerja dan rata-rata lama sekolah. Oleh karena itu, dalam menganalisis hubungan dari indeks pembangunan manusia dengan faktor-faktor tersebut dalam meningkatkan pembangunan dan pertumbuhan ekonomi dilakukan suatu analisis dengan suatu metode dalam statistika. Salah satu metode dalam statistika yang digunakan untuk mengetahui hubungan antara variabel yang mempengaruhi dengan variabel yang dipengaruhi adalah analisis regresi (Maulida, 2022).

Analisis regresi berganda merupakan suatu metode statistik yang digunakan menguji adanya hubungan variabel independen dan variabel dependen. Analisis regresi dalam bentuk linear dengan lebih dari satu variabel independen disebut dengan analisis regresi linear berganda (Kartiningrum, dkk., 2022). Asumsi yang harus dipenuhi dalam analisis regresi adalah tidak adanya hubungan antara variabel independen yang selanjutnya disebut sebagai multikolinearitas (Nugraha, 2022). Penelitian ini menggunakan regresi komponen utama dalam mengatasi masalah multikolinearitas dalam model

regresi, hal tersebut dikarenakan regresi komponen utama menghilangkan korelasi antara variabel independen secara bersih dengan tetap mempertahankan variabel dependen dan dapat digunakan untuk semua data penelitian baik data musiman atau non musiman (Novi, 2016). Oleh karena itu, penelitian ini bertujuan untuk menerapkan regresi komponen utama untuk mengatasi masalah multikolinearitas pada pengaruh produk domestik regional bruto, penduduk miskin, tingkat pengangguran terbuka, jumlah penduduk, angka harapan hidup, jumlah angkatan kerja dan rata-rata lama sekolah terhadap indeks pembangunan manusia.

## 2. METODOLOGI

Palam penelitian ini, peneliti menggunakan pendekatan deskriptif kuantitatif. Alat yang digunakan dalam analisis data dalam penelitian ini adalah *software SPSS* dan *Microsoft Excel* sedangkan data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) mengenai Indeks Pembangunan Manusia (IPM) serta faktor-faktor yang mempengaruhinya.

Variabel-variabel yang terkait dengan penelitian ini adalah sebagai berikut.

Tabel 1.  
Variabel Penelitian

No.	Simbol	Variabel
1	$Y$	Indeks Pembangunan Manusia
2	$X_1$	Produk Domestik Regional Bruto (PDRB)
3	$X_2$	Penduduk Miskin
4	$X_3$	Tingkat Pengangguran Terbuka
5	$X_4$	Jumlag Penduduk
6	$X_5$	Angka Harapan Hidup
7	$X_6$	Jumlah Angkatan Kerja
8	$X_7$	Rata-rata Lama Sekolah

Berikut adalah langkah-langkah dalam penelitian ini.

1. Melakukan standarisasi data penelitian berupa pemusatan dan penskalaan data menggunakan persamaan berikut:

$$Y_i^* = \frac{Y_i - \bar{Y}}{S_Y} \quad (1)$$

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S_{X_j}} \quad (2)$$

2. Menentukan estimasi untuk parameter regresi pada data yang telah di standarisasi menggunakan metode

Ordinary Least Square (OLS) pada persamaan berikut:

$$b = (X'X)^{-1}X'Y \quad (3)$$

selanjutnya ditentukan persamaan regresi linear berganda berdasarkan persamaan:

$$Y = \beta_0 + \beta_1X_1 + \dots + \beta_kX_k + \varepsilon \quad (4)$$

3. Uji mutikolinearitas berdasarkan nilai VIF yang diperoleh dari persamaan berikut:

$$VIF = \frac{1}{1 - R_i^2} \quad (5)$$

Jika terdapat multikolinearitas maka diselesaikan dengan metode regresi komponen utama.

4. Regresi komponen utama dalam menyelesaikan masalah multikolinearitas sebagai berikut.

- a. Menentukan nilai eigen ( $\lambda$ ) dari matriks korelasi ( $\rho$ ) menggunakan persamaan:

$$\det(\rho - \lambda I) = 0 \quad (6)$$

- b. Menentukan vektor eigen atau skor komponen utama ( $K$ ) untuk nilai eigen lebih dari satu menggunakan persamaan:

$$(\rho - \lambda I)a_j = 0 \quad (7)$$

- c. Meregresikan variabel dependen ( $Y$ ) terhadap komponen utama ( $K$ ) menggunakan persamaan

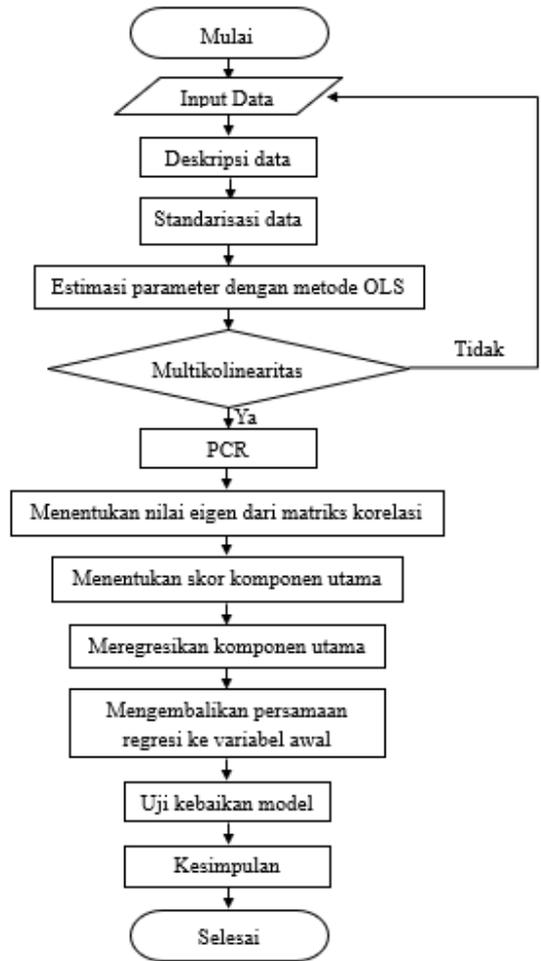
$$Y = \beta_0 + \beta_1K_1 + \beta_2K_2 + \dots + \beta_kK_k + \varepsilon \quad (8)$$

- d. Mentransformasi persamaan regresi linear berganda pada (c) kedalam bentuk variabel standar ( $Z$ ).

- e. Mentransformasi persamaan regresi linear bergana dengan variabel standar ( $Z$ ) ke dalam model regresi linear berganda dengan variabel asal ( $X$ ).

- f. Menguji kebaikan model regresi yang diperoleh.

Adapun prosedur penelitian dapat dilihat pada diagram alir (*flowchart*) berikut:



Gambar 1. Diagram Alir Penelitian

### 3. HASIL DAN PEMBAHASAN

#### Standarisasi Data

Berdasarkan data yang diperoleh dapat diketahui bahwa setiap variabel yang digunakan dalam penelitian ini memiliki satuan yang berbeda sehingga rentang data yang diperoleh berdasarkan selisih nilai rata-rata dan standar deviasi cukup jauh. Oleh karena itu perlu dilakukan standarisasi data menggunakan persamaan (1) dan persamaan (2).

#### Estimasi Parameter Regresi

Data pengamatan yang telah terstandarisasi ditentukan estimator untuk parameter regresinya menggunakan metode *Ordinary Least Square* (OLS). Hasil estimasi untuk parameter regresi dengan metode diperoleh estimator untuk 4 parameter regresi disajikan pada tabel. 1 berikut.

Tabel 1  
Estimasi Parameter Regresi

Variabel	Estimasi $\beta_j$
Konstan	$-3.239 \times 10^{-15}$
$Z_1$	0.144
$Z_2$	0.321
$Z_3$	0.138
$Z_4$	-1.140
$Z_5$	0.360
$Z_6$	1.108
$Z_7$	0.573

Berdasarkan tabel di atas diperoleh model regresi untuk data yang terstandarisasi yaitu:

$$Y = -3.239 \times 10^{-15} + 0.144Z_1 + 0.321Z_2 + 0.138Z_3 - 1.104Z_4 + 0.360Z_5 + 1.108Z_6 + 0.573Z_7$$

### Uji Multikolinieritas

Berdasarkan persamaan (5) diperoleh nilai dari VIF pada variabel independen ( $Z$ ) yang disajikan pada tabel 2 berikut.

Tabel 2.  
Nilai VIF

Variabel	VIF	Keputusan
$Z_1$	1.301	Tidak ada multikolinieritas
$Z_2$	8.106	Tidak ada multikolinieritas
$Z_3$	1.953	Tidak ada multikolinieritas
$Z_4$	147.872	Ada multikolinieritas
$Z_5$	1.342	Tidak ada multikolinieritas
$Z_6$	116.295	Ada multikolinieritas
$Z_7$	2.246	Tidak ada multikolinieritas

Berdasarkan tabel (2) terjadi multikolinieritas pada variabel  $Z_4$  dan  $Z_6$  hal ini dikarenakan nilai VIF yang didapatkan lebih dari atau sama dengan sepuluh. Oleh karena itu, digunakan metode regresi komponen utama untuk mengatasi masalah multikolinieritas pada model regresi.

### Regresi Komponen Utama

#### Matriks Korelasi

Matriks korelasi ( $\rho$ ) dengan komponen matriks tersebut diperoleh menggunakan persamaan:

$$r_{ij} = \frac{n \sum_{i=1}^k \sum_{j=1}^k Z_i Z_j - \sum_{i=1}^k Z_i \sum_{j=1}^k Z_j}{\sqrt{n \sum_{j=1}^k Z_i^2 - (\sum_{i=1}^k Z_i)^2} \sqrt{n \sum_{j=1}^k Z_j^2 - (\sum_{j=1}^k Z_j)^2}} \quad (9)$$

Sehingga diperoleh matriks korelasi ( $\rho$ ) adalah sebagai berikut:

$$\rho = \begin{pmatrix} 1 & -0.111 & 0.388 & -0.103 & 0.254 & -0.102 & 0.405 \\ -0.111 & 1 & -0.111 & 0.893 & -0.028 & 0.863 & -0.355 \\ 0.388 & -0.111 & 1 & 0.05 & 0.324 & 0.055 & 0.634 \\ -0.103 & 0.893 & 0.05 & 1 & 0.052 & 0.994 & -0.185 \\ 0.254 & -0.028 & 0.324 & 0.052 & 1 & 0.86 & 0.398 \\ -0.102 & 0.863 & 0.055 & 0.994 & 0.86 & 1 & -0.159 \\ 0.405 & -0.355 & 0.634 & -0.185 & 0.398 & -0.159 & 1 \end{pmatrix}$$

### Nilai Eigen

Nilai eigen dari matriks korelasi ( $\rho$ ) disajikan pada tabel 4.3 berikut yang diperoleh menggunakan persamaan (6).

Tabel 3.

Komponen	Nilai Eigen
$K_1$	3.03848
$K_2$	2.30896
$K_3$	0.923666
$K_4$	0.675443
$K_5$	-0.355031
$K_6$	0.319234
$K_7$	0.0892509

Setelah diperoleh nilai eigen dari matriks korelasi selanjutnya adalah menghitung keragaman total dari masing-masing komponen. Berikut adalah keragaman total yang diperoleh yang disajikan pada tabel berikut:

Tabel 4.  
Keragaman Total

Komponen	Keragaman Total
$K_1$	0.43407
$K_2$	0.76392
$K_3$	0.895872
$K_4$	0.992364
$K_5$	0.9416454
$K_6$	0.98725
$K_7$	1.000

Berdasarkan tabel 4 terdapat dua komponen yang memiliki nilai eigen lebih dari satu yaitu komponen pertama ( $K_1$ ) dan komponen kedua ( $K_2$ ). Komponen pertama menjelaskan keragaman sebesar 42.687% dari total kumulatif sedangkan komponen kedua mampu menjelaskan 73.295% dari total kumulatif. Dari perolehan nilai kumulatif maka dapat disimpulkan bahwa kedua komponen utama tersebut mampu menjelaskan variabel dependen.

### Vektor Eigen (Skor Komponen Utama)

Berdasarkan nilai eigen untuk komponen  $K_1$  dan  $K_2$  selanjutnya akan dicari vektor eigen (skor komponen

utama) berdasarkan persamaan (7). Vektor eigen (skor komponen utama) disajikan pada tabel berikut.

Tabel 5  
Nilai Skor Komponen Utama

Variabel	$K_1$	$K_2$
$Z_1$	-0.006	0.291
$Z_2$	0.321	-0.032
$Z_3$	0.0409	0.373
$Z_4$	0.347	0.044
$Z_5$	0.061	0.295
$Z_6$	0.345	0.054
$Z_7$	-0.043	0.367

Berdasarkan tabel di atas, diperoleh hubungan antara variabel  $Z$  dengan komponen utama:

$$K_1 = -0.006Z_1 + 0.321Z_2 + 0.049Z_3 + 0.347Z_4 + 0.061Z_5 + 0.345Z_6 - 0.043Z_7$$

$$K_2 = 0.291Z_1 - 0.032Z_2 + 0.373Z_3 + 0.044Z_4 + 0.295Z_5 + 0.054Z_6 + 0.367Z_7$$

### Meregresikan Komponen Utama

Estimasi untuk parameter dari regresi variabel dependen  $Y$  dengan komponen  $K_1$  dan  $K_2$  diperoleh menggunakan metode *Ordinary Least Square* (OLS) pada persamaan (3).

Tabel 6  
Estimasi Parameter Regresi

Variabel	Estimasi $\beta_j$
Konstan	65.742
$K_1$	4.422
$K_2$	0.967

Berdasarkan tabel di atas diperoleh model regresi komponen utama sebagai berikut.

$$Y = 65.742 + 4.422K_1 + 0.967K_2$$

Transformasi regresi komponen utama dengan mensubstitusi persamaan  $K_1$  dan  $K_2$  maka akan diperoleh persamaan regresi dengan variabel independen  $Z$  selanjutnya jika ditransformasi kembali dalam variabel dependen awal  $X$  diperoleh persamaan regresi berikut.

$$Y = 42.548 + 0.00000991X_1 + 0.0371X_2 + 0.343X_3 + 0.000005949X_4 + 0.2532X_5 + 0.000012947X_6 + 0.1348X_7$$

### Uji Keباikaaan Model

Setelah melakukan regresi komponen utama selanjutnya dilakukan uji kebaikan model berdasarkan nilai  $R^2$ . Berdasarkan persamaan

$$R^2 = \frac{JKR}{JKT} = 1 - \frac{JKG}{JKT}$$

diperoleh nilai  $R^2$  sebesar 0,834 artinya model yang diperoleh adalah model yang baik dan dapat menjelaskan 83,4% dari total keragaman data dan dapat juga diinterpretasikan pengaruh dari variabel independen secara simultan terhadap variabel dependen sebesar 83,4% sedangkan 16.6% dipengaruhi oleh variabel lain yang tidak diuji

### 4. KESIMPULAN

Berdasarkan hasil dan pembahasan di atas diperoleh model regresi komponen utama sebagai berikut:

$$Y = 42.548 + 0.00000991X_1 + 0.0371X_2 + 0.343X_3 + 0.000005949X_4 + 0.2532X_5 + 0.000012947X_6 + 0.1348X_7$$

Dengan nilai koefisien determinasi ( $R^2$ ) sebesar 0,834.

### UCAPAN TERIMA KASIH

Judul untuk ucapan terima kasih dan referensi tidak diberi nomor. Terima kasih disampaikan kepada Sekawan Team yang telah meluangkan waktu untuk membuat template ini.

### REFERENSI

- Amelia, S., & Putra, A. A. (2023). Regresi Komponen Utama dalam Mengatasi Multikolinearitas pada Faktor yang Mempengaruhi Pendapatan Asli Daerah di Sumatera Barat. *Jurnal Pendidikan Tambusai*, 7(2), 10906-10914.
- Hadijati, M.(2015). Analisis regresi. FMIPA-Universitas Mataram.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (Sixth Edition). New York: Pearson Education.
- Kartiningrum, E. D., Notobroto, H. B., Otok, B. W., Kumarijati, N. E., & Yuswatiningih, E. (2022). *Aplikasi Regresi Dan Korelasi Dalam Analis Data Hasil Penelitian*. Mojokerto:STIKesMajapahit
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied Linier Statistical Models* (Fifth Edition). McGraw-Hill.
- Marpaung, I., & Winarto, W. (2018). Pengaruh Pengembangan Karir Terhadap Penilaian Prestasi Kerja (Studi Kasus

- Pada Pt. Pln (Persero) Wilayah Sumatera Utara). *Jurnal Ilmiah METHONOMI*, 4(1), 79-86.
- Maulida, Rizka (2022). Perbandingan Principal Component Regression Dan Regresi Ridge Pada Analisis Faktor-Faktor Indeks Pembangunan Manusia (Skripsi, Universitas Islam Negeri Maulana Malik Ibrahim Malang).
- Novi B. P. 2016. Perbandingan Regresi Komponen Utama dengan Regresi Ridge untuk Mengatasi Masalah Multikolinieritas. Skripsi. Semarang: Universitas Negeri Semarang.
- Nugraha, B. (2022). *Pengembangan Uji Statistik: Implementasi Metode Regresi Linear Berganda dengan Pertimbangan Uji Asumsi Klasik*. Sukoharjo: Pradina Pustaka.
- Nurjanah (2023). Efektivitas Analisis Regresi Komponen Utama Robust dengan Metode MCD-LTS pada Data Indeks Pembangunan Manusia (IPM) Provinsi Sumatera Utara. (Skripsi, Universitas Lampung).
- Pratowo, N. I. (2012). Analisis faktor-faktor yang berpengaruh terhadap Indeks Pembangunan Manusia. *Jurnal Studi Ekonomi Indonesia*, 1(1), 15-31.
- Rahmawati, Raehani (2023). Perbandingan Regresi Ridge dan Principal Component Regression dalam Mengatasi Multikolinieritas pada Faktor-Faktor yang Memengaruhi Kemiskinan di Indonesia. (Skripsi, Universitas Mataram).
- Rosyadah, J. A. (2021). Determinan Indeks Pembangunan Manusia (IPM) Provinsi Nusa Tenggara Timur (NTT). *EFFICIENT: Indonesian Journal of Development Economics*, 4(1), 1080-1092.
- Sahir, S.H. (2021). *Metode Penelitian*. Medan : KBM Indonesia.
- Si'lang, I. L. S., Hasid, Z., & Priyagus, P. (2019). Analisis faktor-faktor yang berpengaruh terhadap indeks pembangunan manusia. *Jurnal Manajemen*, 11(2), 159-169.