

**APLIKASI REGRESI PARTIAL LEAST SQUARE UNTUK ANALISIS
HUBUNGAN FAKTOR-FAKTOR YANG MEMPENGARUHI INDEKS
PEMBANGUNAN MANUSIA DI KOTA YOGYAKARTA**

Marwah Masruroh¹, Retno Subekti²

^{1,2} Fakultas MIPA, Universitas Negeri Yogyakarta

e-mail: ¹ mar.marwa.wa27@gmail.com, ² retnosubekti@uny.ac.id

DOI: 10.14710/medstat.9.2.75-85

Abstract

Human Development Index is one of the indicators to measure the success of a region in the field of human development sector. There are several factors that affect Human Development Index, such as life expectancy, the literacy rate, the average length of the school, and the index of purchasing power. The aim in this paper is to analyze the relationship between factors that affect Human Development Index in Yogyakarta using regression analysis. One of the assumptions of classical regression is not going multicollinierity. Multicollinierity cause misinterpretation of regression coefficients with Ordinary Least Square (OLS) method. One method used to overcome multicollinierity is Partial Least Square (PLS). The result of Human Development Index data analysis showed there was a high correlation between the predictor variables or in other words going multicollinierity, so using PLS method, we obtained adjusted R^2 of 99.3% Human Development Index variables can be explained by the four predictor variables. By using PLS method, multicollinierity resolved in the problem of violation in the linear regression assumption.

Keywords: *IPM, OLS, regression, PLS.*

1. PENDAHULUAN

Indeks Pembangunan Manusia (IPM) memperlihatkan tingkat pembangunan sumber daya manusia di suatu wilayah, sehingga dapat dijadikan sebagai tolak ukur angka kesejahteraan suatu daerah atau negara. IPM merupakan indikator gabungan dari beberapa indikator, yaitu indikator kesehatan (ditunjukkan dengan indeks angka harapan hidup waktu lahir), indikator pendidikan (ditunjukkan dengan indeks angka melek huruf dan rata-rata lama sekolah), serta indikator ekonomi (ditunjukkan dengan konsumsi per kapita yang disesuaikan/indeks daya beli penduduk).

Analisis regresi adalah teknik statistika yang dapat digunakan untuk menjelaskan pengaruh variabel prediktor terhadap variabel respon. Model regresi linier dapat diperoleh dengan melakukan estimasi terhadap parameter-parameternya dengan menggunakan metode tertentu. Adapun metode yang dapat digunakan untuk mengestimasi parameter model regresi linier adalah metode *Ordinary Least Square* (OLS). Setelah diperoleh dugaan regresi linier selanjutnya dilakukan pengujian kelayakan, uji parameter, serta uji asumsi klasik (Aryani dan Retno, 2014). Model regresi dikatakan baik jika memenuhi

asumsi klasik, yaitu tidak terjadi autokorelasi, tidak terjadi heteroskedastisitas, data berdistribusi normal serta tidak terjadi multikolinieritas.

Multikolinieritas adalah terjadinya korelasi antar variabel-variabel prediktor yang menyebabkan kesalahan interpretasi koefisien regresi dengan metode OLS. Salah satu metode yang digunakan untuk mengatasi multikolinieritas adalah *Partial Least Square* (PLS) (Abdi dan Williams, 2012). Regresi PLS dapat diperoleh melalui regresi sederhana maupun berganda dengan mengambil kesimpulan dari uji signifikansi untuk memilih variabel prediktor pembangun komponen PLS dan menentukan banyaknya komponen PLS yang terbentuk. Tujuan PLS adalah membentuk komponen yang dapat menangkap informasi dari variabel prediktor untuk memprediksi variabel respon.

2. TINJAUAN PUSTAKA

2.1. Indeks Pembangunan Manusia

Indeks Pembangunan Manusia (IPM) biasa digunakan untuk mengukur keberhasilan atau kinerja suatu wilayah dalam bidang pembangunan manusia. Pembangunan manusia dapat dilihat sebagai pembangunan kemampuan melalui perbaikan taraf kesehatan, pengetahuan, dan ketrampilan, sekaligus pemanfaatan ketrampilan mereka tersebut. Pembangunan manusia sangatlah penting dalam upaya mengurangi tingkat kemiskinan dan kesenjangan, karena pendidikan dan kesehatan yang baik memungkinkan penduduk miskin untuk meningkatkan nilai asetnya (BPS, 2013).

Indeks Pembangunan Manusia (IPM) merupakan indikator gabungan dari beberapa indikator, yaitu indikator kesehatan (ditunjukkan dengan indeks angka harapan hidup waktu lahir), indikator pendidikan (ditunjukkan dengan indeks angka melek huruf dan rata-rata lama sekolah), serta indikator ekonomi (ditunjukkan dengan konsumsi per kapita yang disesuaikan/indeks daya beli penduduk). Ketiga indikator ini dianggap dapat mengukur tingkat kesejahteraan dan keberhasilan pembangunan manusia di suatu wilayah.

2.2. Analisis Regresi

Analisis regresi merupakan studi mengenai ketergantungan variabel respon dengan salah satu atau beberapa variabel prediktor yang umumnya dinyatakan dalam persamaan matematik (Marcus dkk, 2012). Terdapat dua jenis model regresi linier yaitu regresi linier sederhana dan berganda. Model regresi linier sederhana maupun model regresi berganda dapat diperoleh dengan melakukan estimasi terhadap parameter-parameternya dengan menggunakan metode tertentu. Adapun metode yang dapat digunakan untuk mengestimasi parameter model regresi adalah metode *Ordinary Least Square* (OLS).

Setelah didapatkan dugaan regresi linier langkah selanjutnya adalah pengujian kelayakan dan uji parameter model regresi. Kelayakan model regresi dapat dinilai dari nilai koefisien determinasi (R^2), nilai koefisien determinasi yang tinggi (mendekati 1) menginterpretasikan bahwa model regresi yang diperoleh baik. Pengujian parameter dapat diukur melalui uji signifikansi F dan uji signifikansi t . Jika nilai signifikansi F lebih kecil dari 0.05 (α), maka disimpulkan variabel-variabel prediktor secara bersama-sama berpengaruh terhadap variabel respon, dan jika signifikansi t lebih kecil dari 0.05 (α), maka disimpulkan variabel-variabel prediktor secara parsial berpengaruh terhadap variabel respon. Dalam statistika, sebuah model regresi dikatakan baik atau cocok jika dipenuhi

asumsi-asumsi ideal (klasik), yakni normalitas, tidak ada autokorelasi, serta tidak terjadi heteroskedastisitas dan multikolinieritas (Andryan, 2010).

2.3. Partial Least Square (PLS)

Regresi *Partial Least Square* (PLS) univariat adalah sebuah model yang menghubungkan antara sebuah variabel respon Y dengan sekumpulan variabel prediktor X . Regresi PLS merupakan salah satu metode untuk mengatasi multikolinieritas yang dapat diperoleh melalui regresi sederhana maupun berganda dengan mengambil kesimpulan dari uji signifikansi. Uji signifikansi bertujuan untuk memilih variabel prediktor pembangun komponen PLS dan menentukan banyaknya komponen PLS yang terbentuk. Tujuan PLS adalah membentuk komponen yang dapat menangkap informasi dari variabel prediktor untuk memprediksi variabel respon (Bastien *et al*, 2004). Model regresi PLS dengan m komponen dapat dituliskan sebagai berikut:

$$Y = \sum_{h=1}^m c_h t_h + \varepsilon \quad (1)$$

dengan:

Y : variabel respon

c_h : koefisien regresi Y terhadap t_h

$t_h = \sum_{j=1}^p w_{(h)j} X_j$: komponen utama ke- h yang tidak saling berkorelasi ($h = 1, 2, \dots, m$)

Dengan syarat komponen PLS $t_h = \sum_{j=1}^p w_{(h)j} X_j$ orthogonal, sehingga parameter c_h dan w_h dalam Persamaan (1) dapat diestimasi.

1. Perhitungan Komponen PLS Pertama, t_1

Komponen PLS pertama (t_1) adalah kombinasi linier dari variabel prediktor X_j dengan koefisien pembobot w_1 . Komponen pertama $t_1 = Xw_1$ didefinisikan sebagai berikut:

$$t_1 = \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(x_j, y)^2}} \sum_{j=1}^p \text{cor}(x_j, y) x_j^* \quad (2)$$

dengan:

x_j^* : x_j terstandarisasi

$\text{cor}(x_j, y)$: korelasi variabel x_j dengan y

2. Perhitungan Komponen PLS Kedua, t_2

Komponen PLS kedua didapatkan dengan melakukan regresi sederhana y terhadap t_1 dan masing-masing x_j terlebih dahulu kemudian regresi antara x_j terhadap t_1 . Variabel-variabel x_j yang digunakan hanya variabel yang berkontribusi secara signifikan dalam menjelaskan y pada t_1 . Komponen PLS kedua dapat dituliskan sebagai berikut:

$$t_2 = \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(y, x_j | t_1)^2}} \sum_{j=1}^p \text{cor}(y, x_j | t_1) x_{1j}^* \quad (3)$$

dengan x_{1j}^* adalah residu yang telah distandarisasi dan dihasilkan dari regresi x_j terhadap t_1 .

3. Perhitungan Komponen PLS ke- h , t_h
Seperti langkah sebelumnya, variabel yang digunakan adalah variabel-variabel yang signifikan dalam menjelaskan y pada $t_1, t_2, \dots, t_{h-1}$. Komponen PLS ke- h didefinisikan sebagai berikut:

$$t_h = \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(x_{(h-1)j}, y)^2}} \sum_{j=1}^p \text{cor}(x_{(h-1)j}, y) x_{(h-1)j}^* \quad (4)$$

dengan $x_{(h-1)j}^*$ adalah residu standar dari regresi setiap x_j terhadap $t_1, t_2, \dots, t_{h-1}$. Perhitungan komponen PLS berhenti ketika tidak ada lagi variabel prediktor yang signifikan membangun komponen PLS.

3. METODE PENELITIAN

3.1. Sumber Data dan Variabel Penelitian

Penelitian ini akan membahas faktor-faktor yang mempengaruhi IPM di Kota Yogyakarta dengan menggunakan metode PLS. Data yang digunakan dalam penelitian ini adalah data hasil Survei Sosial Ekonoi Nasional (SUSENAS) di Kota Yogyakarta oleh Badan Pusat Statistik (BPS). Variabel yang digunakan pada penelitian ini sebanyak 5 variabel. Variabel tersebut terbagi atas satu variabel respon dan empat variabel prediktor. Variabel yang berperan sebagai variabel respon (Y) adalah Indeks Pembangunan Manusia (IPM), sedangkan variabel prediktor yang dipakai penelitian ini meliputi:

1. Angka harapan hidup (AHH) sebagai x_1
2. Angka melek huruf (AMH) sebagai x_2
3. Rata-rata lama sekolah (MYS) sebagai x_3
4. Indeks Daya Beli (PPP) sebagai x_4

3.2. Metode Analisis

Dalam penelitian ini *software* yang digunakan adalah SPSS. Adapun langkah analisis yang digunakan dalam penelitian ini adalah:

1. Melakukan pemodelan regresi linier berganda dengan metode *Ordinary Least Square* (OLS).
2. Melakukan uji kelayakan, uji parameter model regresi dan uji asumsi klasik.
Kelayakan model regresi dapat dinilai dari nilai koefisien determinasi (R^2), sedangkan pengujian parameter dilakukan secara bersama melalui uji signifikansi F dan pengujian parameter secara parsial melalui uji signifikansi t . Uji asumsi klasik dapat dianalisis dengan *software* SPSS, beberapa diantaranya yaitu: uji non-parametrik Kolmogorov Smirnov untuk uji normalitas, uji *Run Test* untuk uji autokorelasi, uji Glejser untuk uji

heteroskedastisitas, serta melihat nilai *Variance Inflation Factor* (VIF) dan *Tolerance* (TOL) untuk uji multikolinieritas.

3. Melakukan pemodelan regresi dengan metode PLS untuk mengatasi multikolinieritas. Dalam pembentukan komponen PLS, digunakan variabel respon Y yang distandarisasi (y^*) dan variabel-variabel prediktor yang terpusat. Langkah pertama dalam pembentukan model regresi *partial least square* adalah dengan mencari banyaknya komponen PLS yang tepat sehingga didapatkan model regresi yang bebas multikolinieritas. Sebelum pembentukan komponen PLS, terlebih dahulu dilakukan regresi Y yang distandarisasi (y^*) terhadap $t_1, t_2, \dots, t_{h-1}$ dan terhadap masing-masing variabel prediktor terpusat. Variabel-variabel prediktor yang digunakan hanya variabel yang signifikan dalam menjelaskan y pada $t_1, t_2, \dots, t_{h-1}$, sehingga komponen PLS berhenti jika sudah tidak ada lagi variabel prediktor yang signifikan membangun komponen PLS. Komponen PLS ke- h dapat dihitung menggunakan Persamaan (4). Setelah didapatkan komponen PLS yang sesuai, model regresi *Partial Least Square* dengan m komponen dapat dibentuk dan ditulis menggunakan Persamaan (1).
4. Tranformasi Komponen PLS ke variabel asli t_h
Persamaan (1) yang diperoleh setelah perhitungan kemudian ditulis ke dalam bentuk variabel aslinya, yaitu:

$$\begin{aligned}
 Y &= \sum_{h=1}^m c_h t_h + \varepsilon \\
 &= \sum_{h=1}^m c_h \left(\sum_{j=1}^p w_{(h)j} X_j \right) + \varepsilon \\
 &= \sum_{j=1}^p \sum_{h=1}^m c_h w_{(h)j} X_j + \varepsilon \\
 &= \sum_{j=1}^p \left(\sum_{h=1}^m c_h w_{(h)j} \right) X_j + \varepsilon \\
 Y &= \sum_{j=1}^p b_j X_j + \varepsilon
 \end{aligned}$$

dengan:

Y : variabel respon

c_h : koefisien regresi Y terhadap t_h

X_j : matriks variabel prediktor

$t_h = \sum_{j=1}^p w_{(h)j} X_j$: komponen utama ke- h yang tidak saling berkorelasi ($h = 1, 2, \dots, m$)

$b_j = \sum_{h=1}^p c_h w_{(h)j}$: vektor koefisien regresi Y terhadap X_j

$w_{(h)j}$: koefisien bobot untuk variabel X_j pada komponen utama PL ke- h

ε : vektor error

4. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data IPM Kota Yogyakarta dan faktor-faktor yang mempengaruhinya dari tahun 2005 sampai dengan tahun 2013. Data yang digunakan diberikan pada Tabel 1.

Tabel 1. Data IPM Kota Yogyakarta Periode 2005-2013

Tahun	IPM	AHH	AMH	MYS	PPP
2005	77.7	72.9	97.1	10.8	639.1
2006	77.8	73.1	97.1	10.8	639.2
2007	78.1	73.1	97.7	10.9	640.6
2008	78.9	73.2	97.7	11.4	645.1
2009	79.3	73.4	98.0	11.5	647.6
2010	79.5	73.4	98.0	11.5	649.7
2011	79.8	73.5	98.0	11.5	653.8
2012	80.2	73.5	98.1	11.6	657.6
2013	80.5	73.7	98.4	11.6	658.8

Dari data indeks IPM pada Tabel 1 dilakukan regresi linier dengan metode OLS dengan menggunakan program SPSS. Analisis regresi ini bertujuan untuk mengetahui hubungan variabel-variabel prediktor terhadap variabel respon, sehingga diperoleh hasil persamaan regresi linier dugaannya adalah sebagai berikut:

$$\hat{Y} = -36.683 + 0.508 x_1 + 0.22 x_2 + 0.777 x_3 + 0.074 x_4 \quad (5)$$

Nilai *adjusted R²* sebesar 0.999, yang berarti sebesar 99.9% variabel IPM dapat dijelaskan oleh keempat variabel prediktor. Nilai signifikansi F sebesar 0.000 (kurang dari 0.05) yang berarti variabel prediktor secara bersama-sama berpengaruh terhadap variabel respon. Nilai signifikansi *t* untuk keempat variabel prediktor berturut-turut sebesar 0.034, 0.036, 0.001 dan 0.000 dimana semuanya kurang dari 0.05 sehingga disimpulkan variabel prediktor secara parsial berpengaruh terhadap variabel respon.

Terdapat empat asumsi klasik pada regresi yang harus dipenuhi untuk menyatakan bahwa model regresi tersebut baik, yaitu tidak terjadi autokorelasi, tidak terjadi heteroskedastisitas, data berdistribusi normal serta tidak terjadi multikolinieritas. Hasil analisa data IPM menyatakan bahwa semua asumsi klasik terpenuhi kecuali multikolinieritas. Hal ini ditunjukkan dengan adanya korelasi yang tinggi antar variabel yang dapat dilihat pada Tabel 2.

Korelasi yang tinggi antar variabel pada Tabel 2 mengidentifikasi adanya multikolinieritas, sehingga akan digunakan metode PLS untuk mengatasinya. Sebelum pembentukan komponen pertama PLS, terlebih dahulu dilakukan regresi *y* yang distandarisasi (*y*^{*}) terhadap masing-masing *x_j* terpusat untuk mengetahui variabel-variabel manakah yang signifikan membangun komponen PLS pertama. Hasil signifikansi variabel *x_j* masing-masing sebesar 0.000. Uji signifikansi koefisien regresi tersebut menunjukkan bahwa dengan taraf nyata 5% semua variabel signifikan membangun komponen PLS pertama karena nilai sig < 0.05. Perhitungan komponen PLS pertama, *t₁* adalah sebagai berikut:

$$t_1 = \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(x_j, y)^2}} \sum_{j=1}^p \text{cor}(x_j, y) x_j^*$$

$$t_1 = \frac{0.969x_1^* + 0.939x_2^* + 0.952x_3^* + 0.986x_4^*}{\sqrt{0.969^2 + 0.939^2 + 0.952^2 + 0.986^2}}$$

$$= 0.5038x_1^* + 0.4882x_2^* + 0.4950x_3^* + 0.5127x_4^* \quad (6)$$

Tabel 2. Korelasi antar Variabel

		IPM	AHH	AMH	MYS	PPP
IPM	Pearson Correlation	1	0.969**	0.939**	0.952**	0.986**
	Sig. (2-tailed)		0.000	0.000	0.000	0.000
	N	9	9	9	9	9
AHH	Pearson Correlation	0.969**	1	0.930**	0.899**	0.950**
	Sig. (2-tailed)	0.000		0.000	0.001	0.000
	N	9	9	9	9	9
AMH	Pearson Correlation	0.939**	0.930**	1	0.908**	0.895**
	Sig. (2-tailed)	0.000	0.000		0.001	0.001
	N	9	9	9	9	9
MYS	Pearson Correlation	0.952**	0.899**	0.908**	1	0.897**
	Sig. (2-tailed)	0.000	0.001	0.001		0.001
	N	9	9	9	9	9
PPP	Pearson Correlation	0.986**	0.950**	0.895**	0.897**	1
	Sig. (2-tailed)	0.000	0.000	0.001	0.001	
	N	9	9	9	9	9

**, Correlation is significant at the 0.01 level (2-tailed).

Sebelum pembentukan komponen PLS kedua, terlebih dahulu diperiksa apakah komponen kedua ini masih diperlukan. Hal tersebut dilakukan dengan cara meregresikan antara y yang distandarisasi (y^*) terhadap t_1 dan masing-masing variabel x_j terpusat. Hasil signifikansi variabel x_j (AHH, AMH, MYS dan PPP) berturut-turut sebesar 0.725, 0.032, 0.972 dan 0.002. Variabel yang signifikan membangun komponen PLS kedua adalah AMH (x_2) dan PPP (x_4), sehingga akan dihitung komponen PLS kedua. Untuk membangun komponen PLS kedua diperlukan koefisien residu x_{12} yaitu residu yang dihasilkan dari persamaan regresi antara x_2 terhadap t_1 dan residu x_{14} yaitu residu yang dihasilkan dari persamaan regresi antara x_4 terhadap t_1 , yaitu:

$$x_2 = p_{12}t_1 + x_{12}$$

$$x_4 = p_{14}t_1 + x_{14}$$

p_{12} adalah koefisien t_1 pada regresi x_2 terhadap t_1 dan p_{14} adalah koefisien t_1 pada regresi x_4 terhadap t_1 . Perhitungan komponen PLS kedua, t_2 adalah sebagai berikut:

$$t_2 = \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(y, x_{1j})^2}} \sum_{j=1}^p \text{cor}(y, x_{1j}) x_{1j}^*$$

$$\begin{aligned}
t_2 &= \frac{-0.077x_{12}^* + 0.093x_{14}^*}{\sqrt{(-0.077)^2 + 0.093^2}} \\
&= \frac{-0.638x_{12}}{\sqrt{\text{var}(x_{12})}} + \frac{0.770x_{14}}{\sqrt{\text{var}(x_{14})}} \\
&= \frac{-0.638(x_2 - p_{12}t_1)}{\sqrt{\text{var}(x_{12})}} + \frac{0.770(x_4 - p_{14}t_1)}{\sqrt{\text{var}(x_{14})}} \\
&= \frac{-0.638(x_2 - 0.221t_1)}{0.1169} + \frac{0.770(x_4 - 3.820t_1)}{1.8969} \\
&= -5.458 \frac{x_2}{\sqrt{\text{var}(x_2)}} + 0.406 \frac{x_4}{\sqrt{\text{var}(x_4)}} - 0.663t_1 \\
t_2 &= -0.174x_1^* - 12.494x_2^* - 0.171x_3^* - 0.124x_4^* \tag{7}
\end{aligned}$$

Komponen PLS ketiga dibentuk setelah terlebih dahulu diperiksa apakah komponen ketiga ini masih diperlukan atau tidak, yaitu dengan cara meregresikan y yang telah distandarisasi (y^*) terhadap t_1 , t_2 dan masing-masing variabel x_j terpusat. Hasil signifikansi variabel x_j (AHH, MYS dan PPP) berturut-turut sebesar 0.257, 0.572 dan 0.015. Variabel AMH berkorelasi tinggi dengan t_2 sehingga AMH dikeluarkan dan dianggap tidak signifikan membangun komponen PLS ketiga. Variabel yang signifikan membangun komponen PLS ketiga adalah PPP (x_4), sehingga akan dihitung komponen PLS ketiga. Untuk membangun komponen PLS ketiga diperlukan koefisien residu x_{24} , yaitu residu yang dihasilkan dari persamaan regresi antara x_4 terhadap t_1 dan t_2 , yaitu:

$$x_4 = p_{14}t_1 + p_{24}t_2 + x_{24}$$

p_{14} adalah koefisien t_1 pada regresi x_4 terhadap t_1 dan t_2 dan p_{24} adalah koefisien t_2 pada regresi x_4 terhadap t_1 dan t_2 . Perhitungan komponen PLS ketiga, t_3 adalah sebagai berikut:

$$\begin{aligned}
t_3 &= \frac{1}{\sqrt{\sum_{j=1}^p \text{cor}(x_{2j}, y)^2}} \sum_{j=1}^p \text{cor}(x_{2j}, y)x_{2j}^* \\
t_3 &= \frac{0.058x_{24}^*}{\sqrt{0.058^2}} = x_{24}^* \\
&= \frac{x_{24}}{\sqrt{\text{var}(x_{24})}} \\
&= \frac{x_4 - p_{14}t_1 - p_{24}t_2}{\sqrt{\text{var}(x_{24})}} \\
&= \frac{x_4 - 6.049t_1 - 0.345t_2}{1.5279} \\
&= 0.655 \frac{x_4}{\sqrt{\text{var}(x_4)}} - 3.959t_1 - 0.226t_2
\end{aligned}$$

$$t_3 = -1.955x_1^* + 0.891x_2^* - 1.921x_3^* - 1.916x_4^* \quad (8)$$

Komponen PLS keempat dibentuk setelah terlebih dahulu diperiksa apakah komponen keempat ini masih diperlukan atau tidak, yaitu dengan cara meregresikan y yang telah distandarisasi (y^*) terhadap t_1 , t_2 , t_3 dan masing-masing variabel x_j terpusat. Pada tahap ini ternyata semua variabel prediktor berkorelasi tinggi dengan t_1 , sehingga variabel AHH, AMH, MYS dan PPP dikeluarkan dan dianggap tidak signifikan membangun komponen PLS keempat. Karena semua variabel prediktor tidak ada yang signifikan membangun komponen PLS keempat, maka perhitungan berhenti pada komponen PLS ketiga dan diperoleh komponen baru yaitu t_1 , t_2 , dan t_3 . Kemudian variabel respon y diregresikan terhadap komponen tersebut. Karena variabel respon y berkorelasi tinggi dengan t_1 , maka variabel t_1 dikeluarkan dari model. Sehingga hanya dilakukan regresi antara variabel IPM (y) dengan variabel t_2 dan t_3 , dengan program SPSS diperoleh hasil regresi linier dugaan seperti berikut:

$$\begin{aligned} \hat{Y}_{PLS} &= 79,089 - 0,015 t_2 - 0,177 t_3 \\ &= 79.089 - 0.015(-0.174x_1^* - 12.494x_2^* - 0.171x_3^* - 0.124x_4^*) \\ &\quad - 0.177(-1.955x_1^* + 0.891x_2^* - 1.921x_3^* - 1.916x_4^*) \\ &= 79.089 + 0.349x_1^* + 0.029x_2^* + 0.342x_3^* + 0.341x_4^* \\ &= 79.089 + 0.349\left(\frac{x_1 - \mu_1}{sd_1}\right) + 0.029\left(\frac{x_2 - \mu_2}{sd_2}\right) + 0.342\left(\frac{x_3 - \mu_3}{sd_3}\right) + 0.341\left(\frac{x_4 - \mu_4}{sd_4}\right) \\ &= 79.089 + \frac{0.349}{sd_1}x_1 - \frac{0.349}{sd_1}\mu_1 + \frac{0.029}{sd_2}x_2 - \frac{0.029}{sd_2}\mu_2 + \frac{0.342}{sd_3}x_3 - \frac{0.342}{sd_3}\mu_3 \\ &\quad + \frac{0.341}{sd_4}x_4 - \frac{0.341}{sd_4}\mu_4 \\ &= 79.089 + \frac{0.349}{0.252}x_1 - \frac{0.349(73.311)}{0.252} + \frac{0.029}{0.443}x_2 - \frac{0.029(97.789)}{0.443} \\ &\quad + \frac{0.342}{0.348}x_3 - \frac{0.342(11.289)}{0.348} + \frac{0.341}{7.626}x_4 - \frac{0.341(674.94)}{7.626} \\ &= 79.089 - 147.9143 + 1.3838x_1 + 0.0655x_2 + 0.9827x_3 + 0.0447x_4 \\ \hat{Y}_{PLS} &= -68.8253 + 1.3838AHH + 0.0655AMH + 0.9827MYS + 0.0447PPP \quad (9) \end{aligned}$$

Setelah diperoleh model regresi linier dugaan dengan metode PLS, maka langkah selanjutnya adalah dilakukan uji asumsi klasik, uji parameter dan uji kelayakan model regresi. Dengan menggunakan *software* SPSS diperoleh hasil uji asumsi klasik, dimana semua asumsi klasik pada data IPM terpenuhi, masalah multikolinieritas yang terjadi sebelumnya sudah dapat diatasi dengan metode PLS. pengujian parameter dilakukan secara bersama melalui uji signifikansi F dan uji signifikansi t. Nilai signifikansi F sebesar 0.000 yang berarti kedua komponen PLS secara bersama berpengaruh terhadap variabel respon,

sedangkan nilai signifikansi t untuk kedua variabel prediktor berturut-turut sebesar 0.042 dan 0.000, nilai keduanya masing-masing kurang dari 0.05 sehingga disimpulkan variabel prediktor secara parsial berpengaruh terhadap variabel respon. Pengujian kelayakan dapat dilihat dari nilai $adjusted R^2$. Dari persamaan regresi dugaan dengan metode PLS didapatkan $adjusted R^2$ sebesar 0.993, yang berarti sebesar 99.3% variabel IPM dapat dijelaskan oleh keempat variabel prediktor, sedangkan 0.7% sisanya dijelaskan oleh variabel lainnya.

Meningkatnya nilai AHH, AMH, MYS dan PPP dapat meningkatkan nilai IPM. Koefisien regresi variabel angka harapan hidup (AHH) sebesar 1.3838 memiliki arti jika angka harapan hidup mengalami peningkatan 1% dan rata-rata lama sekolah, angka melek huruf dan indeks daya beli konstan maka IPM meningkat sebesar 1.3838. Kemudian jika angka melek huruf (AMH) mengalami peningkatan 1% dan angka harapan hidup, rata-rata lama sekolah dan indeks daya beli konstan maka IPM meningkat sebesar 0.0655. Jika rata-rata lama sekolah (MYS) meningkat 1% dan angka harapan hidup, angka melek huruf serta indeks daya beli konstan maka IPM meningkat sebesar 0.9827. Jika indeks daya beli (PPP) meningkat 1% dan angka harapan hidup, rata-rata lama sekolah serta angka melek huruf konstan maka IPM meningkat sebesar 0.0447.

5. KESIMPULAN

Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa dengan menggunakan metode PLS, keempat faktor yaitu AHH, AMH, MYS dan PPP berpengaruh positif terhadap IPM. Nilai $adjusted R^2$ sebesar 0.993, hal ini menunjukkan bahwa sebesar 99.3% variabel IPM dapat dijelaskan oleh keempat variabel prediktor.

DAFTAR PUSTAKA

- Abdi, H. dan Williams, L., 2012, *Partial Least Squares Correlation and Partial Least Squares Regression*, Chapter. 23, pp 549-579.
- Andryan, S., 2010, *Uji Asumsi Klasik Dengan SPSS 16.0*, Fakultas Ekonomi UNNES, Semarang.
- Aryani, D. A. dan Retno, S., 2014, *Partial Least Square (PLS) dan Principal Component Regression (PCR) untuk Regresi Linier dengan Multikolinieritas pada Kasus Indeks Pembangunan Manusia di Kabupaten Gunung Kidul*, Skripsi Fakultas MIPA UNY, Yogyakarta, Tidak dipublikasikan.
- Bastien, P., Vnzi, V. dan Tenenhaus, M., 2004, Partial Least Square Generalized Linier Regression, *Computational Statistics & Data Analysis* 48, pp 17-46.
- BPS, 2013, *Indeks Pembangunan Manusia Kota Yogyakarta 2013*, BPS Kota Yogyakarta.
- Marcus, G., Wattimanel, H. dan Lenussa, Y., 2012, Analisis Regresi Komponen Utama untuk Mengatasi Multikolinieritas dalam Analisis Regresi Linier Berganda, *Jurnal Ilmu Matematika dan Terapan*, Vol. 6, No. 1, pp 31-40.