Analysis of Human Development Index in Sumatera Barat Province using Biplot Method

By
Eri Mardison
Badan Pusat Statistik, Sumatra Utara
Corresponding Autor: erim@bps.go.id

ABSTRACT

The Human Development Index (HDI) is calculated from four variables, however, the HDI value cannot directly show which variables are superior in an area and which are not. For this reason, this study aims to analyze the strength of the HDI variables in each region in Sumatera Barat Province, using Biplot Analysis. Important findings resulted from biplot analysis indicate that the Sumatera Barat Province is categorized into five regional groups that are affected by the proximity value of its constituent variables. Out of the five groups, two groups are in an area not adjacent to the variable. The other two groups are in the variable area. While the last group is an extreme region compared to other regional groups. Areas that are not in the area of variables are weak against variables, while the area falling in the variable area is superior in the relevant variable. For extreme groups, Kota Padang is very superior in the adjusted per capita expenditure variable and superior in other variables too. Meanwhile, the Mentawai Islands Regency is very weak in all variables. Kota Padang Panjang is very superior in all variables of education but weak in the economic variable. The Biplot Analysis also revealed that Sijunjung Regency had a fairly good in expenditure, but it was not good enough in other dimensions of the HDI. Things like this are not given enough attention to ordinary HDI analysis, this is what makes the Biplot analysis important to do.

Keywords: Biplot Analysis, HDI, Sumatera Barat.

ABSTRAK

Indeks Pembangunan Manusia (IPM) dihasilkan dari empat komponen, namun demikian nilai IPM tersebut tidak dapat menunjukkan secara langsung komponen mana yang unggul pada suatu daerah dan mana yang tidak. Untuk itu, maka penelitian ini bertujuan untuk menganalisis kekuatan komponen IPM di setiap wilayah di Provinsi Sumatera Barat, dengan menggunakan Analisis Biplot. Temuan penting yang dihasilkan dari analisis biplot menunjukkan bahwa Provinsi Sumatera Barat dikelompokkan ke dalam lima group wilayah yang dipengaruhi oleh nilai kedekatan komponen penyusunnya. Dari lima group tersebut, dua grup berada pada daerah yang tidak berdekatan dengan peubah. Dua grup lainnya berada pada daerah peubah. Sedangkan satu grup terakhir merupakan daerah yang tergolong ekstrim dibandingkan dengan kelompok daerah lainnya. Daerah yang tidak berada pada area peubah, bersifat lemah terhadap peubah. Sementara daerah yang berada pada aera peubah akan unggul pada peubah bersangkutan. Untuk grup

ekstrim, Kota Padang sangat unggul pada komponen pengeluaran per kapita yang disesuaikan dan juga unggul pada komponen-komponen lainnya. Sedangkan Kabupaten Kepulauan Mentawai sangat lemah pada semua komponen. Kota Padang Panjang sangat unggul pada seluruh komponen pendidikan, namun lemah pada komponen ekonomi. Biplot analisis juga mengungkap bahwa Kabupten Sijunjung mempunyai pengeluaran yang cukup baik, tetaPI tidak cukup baik pada dimensi IPM yang lain. Hal-hal seperti ini kurang diperhatikan pada analisis IPM biasa, inilah yang membuat Biplot analisis menjadi penting untuk dilaksanakan

Kata Kunci: *Analisis BIplot, IPM, Sumatera Barat.*

INTRODUCTION

The Goals of all development are humans, all for humans. The Human Development Index (HDI) is one of the indicators to measure the success of the development, according to Junian et al. (2018). Pieces of research have been done by pairing the HDI with other variables such as poverty, population, and economic growth, one of which was done by Prasetyoningrum & Sukmawati (2018). However, not many studies mention and discuss previous studies that do not examine the HDI variable itself. Indonesia began to calculate HDI from 1990. Indonesian HDI increased from 53.2 in 1990 to 71.39 in 2018. It always increased with a different magnitude, Indonesian HDI had only experienced a decline in 1998 due to economic crisis. HDI is one of important indicator to measure the success of human development efforts. In connection with that IPM also became one of the allocators for the General Allocation Fund (DAU). Sumatera Barat Province has a relatively stable of HDI ranking. In 1996 the Province of Sumatra Barat was ranked 11th with an HDI score of 69.2, in 1999 it had been ranked 9th and had risen in 2002 to rank 8th but since 2004 until now it has survived at a number 9th. A region with sufficient stability is interesting to be used as research material. Moreover this research is the first study by raising the main variables of HDI in Sumatera Barat Province.

The new method of HDI consists of three dimensions; the health dimension is represented by the Life Expectancy (LE), the education dimension is represented by Expected Years of School (EYS) and Mean Years of School (MYS), and the economic dimension is represented by adjusted per capita expenditures, as defined by BPS (2019).

Usually, the classification of HDI is low, medium, high, and very high, based on BPS (2019). This research tried to explore the HDI grouping using biplot analysis. Biplot analysis is classified as a double variable exploration analysis, which intended to reveal multiple variable data in a two-dimensional scatter map, so that the data behavior of all involved variables is easy to be interpreted, as stated by Tinungki (2018).

METHODS

This study chose a biplot analysis method since the method can describe the data in a two-dimensional graph (Hatidja, Momuat, 2017). A multivariate method that uses graphs in the form of rows and columns is the analysis biplot. Thus, the objects and variables under study can be displayed. As a double dimensional descriptive statistical technique, the biplot can present a group of observations and variables simultaneously on one image. The characteristics of the variable and the object of observation and the relative position between the object of observation and the variable can be analysed. These advantages and uses are very much in accordance to the research objectives, one of which is to show the superiority of an area and its proximity to other variables. There are four important things that can be obtained on the biplot display, (Leleury & Wokanubun, 2015):
1. Relationships between observed objects.
2. Variability of variables.
3. Relations between variables.

4. Relationships between objects and variables.

Biplot analysis or also called Singular Value composition (SVD) is based on the breakdown of singular values. Biplot can be built from a data matrix, where each column represents a variable and each row represents the object of research.

The Biplot method yield information about the relative position of the original variable. Furthermore, the combination of the original variable plot with the observation plot through superimpose will produce information about the relationship between the variable and the observation.

Based on matrix:

$$_n\mathbf{X}_p = \begin{pmatrix} x_{11} & \cdots & x_{1i} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{k1} & \cdots & x_{ki} & \cdots & x_{kp} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{ni} & \cdots & x_{np} \end{pmatrix}$$

It can generate matrix $\mathbf{G}$ and $\mathbf{H}$ as follows:

$$\mathbf{G} = \begin{pmatrix} g_{11} & g_{12} \\ \vdots & \vdots \\ g_{k1} & g_{k2} \\ \vdots & \vdots \\ g_{n1} & g_{n2} \end{pmatrix} = \begin{pmatrix} \mathbf{g}_1^T \\ \vdots \\ \mathbf{g}_k^T \\ \vdots \\ \mathbf{g}_n^T \end{pmatrix}$$

$$H = \begin{pmatrix} h_{11} & h_{12} \\ \vdots & \vdots \\ h_{i1} & h_{i2} \\ \vdots & \vdots \\ h_{p1} & h_{p2} \end{pmatrix} = \begin{pmatrix} \mathbf{h}_1^T \\ \vdots \\ \mathbf{h}_i^T \\ \vdots \\ \mathbf{h}_p^T \end{pmatrix}$$

The target:

$$\mathbf{g}_k^T = (g_{k1} \quad g_{k2}) \quad represents \ \mathbf{x}_k^T = (x_{k1} \ldots \quad x_{k2} \quad \ldots \quad x_{kp})$$

$$\mathbf{h}_k^T = (h_{i1} \quad h_{k2}) \quad represents \ \mathbf{x}_i^T = (x_{1i} \ldots \quad x_{ki} \quad \ldots \quad x_{ni})$$

Suppose the $_n\mathbf{Y}_p$ matrix is a data matrix and $_n\mathbf{X}_p$ is data matrix that has been corrected for its median, viz. $\mathbf{X} = \mathbf{Y} - \mathbf{(JY)}/n$, where J is a one-element matrix and has the size of nxn. With singular value decomposition obtained::

$$_n\mathbf{X}_p = {}_n\mathbf{U}_r \ \mathbf{L}_r \ \mathbf{A}_p \qquad\qquad (1)$$

where :

  $\mathbf{U}$ and $\mathbf{A}$ are matrices with orthonormal columns ($\mathbf{U}^T\mathbf{U} = \mathbf{A}^T\mathbf{A} = {}_r\mathbf{I}_r$)

  $\mathbf{L}$ is a diagonal matrix with diagonal elements in the form of eigen values

The above equation can be written as:)

$$\mathbf{X} = \mathbf{U} \ \mathbf{L}^\alpha \ \mathbf{L}^{1-\alpha} \mathbf{A}' \qquad\qquad (2)$$

$$\mathbf{X} = {}_n\mathbf{G}_r \ {}_r\mathbf{H}_p$$

By defining $\mathbf{G} = \mathbf{U} \ \mathbf{L}^\alpha$ dan $\mathbf{H} = \mathbf{L}^{1\alpha}\mathbf{A}'$

In special case, where α = 0, then G = U, and H = AL

RESULTS AND DISCUSSIONS

One of the prerequisites for biplot analysis is the composition of data that is normally distributed, (Sholihin et all, 2014). This step is carried out in anticipation of the data transformation process. From the data normality test output using the Kolmogorov-Smirnov test, it appears that the sig value for all variables has a greater value than α = 0.05. So it can be said that the data is normally distributed. Because the data is normally distributed, the transformation process is not necessary.

Table 1. The Result of One-Sample Kolmogorov-Smirnov Test

		LE	EYS	MYS	A. Expenditure
N		19	19	19	19
Normal Parameters[a,b]	Mean	69.7573684 21052620	13.6542	9.0037	10515.11
	Std. Deviation	2.77905427 9797362	1.07918	1.51886	1911.860
Most Extreme Differences	Absolute	.108	.155	.224	.147
	Positive	.077	.155	.224	.147
	Negative	-.108	-.113	-.130	-.106
Test Statistic		.108	.155	.224	.147
Asymp. Sig. (2-tailed)		.200[c,d]	.200[c,d]	.013[c]	.200[c,d]

a. Test distribution is Normal.

b. Calculated from data.

c. Lilliefors Significance Correction.

d. This is a lower bound of the true significance.

Tabel 2. The Result of Communality

	Initial	Extraction
Life Expectancy	1,000	0,866
Expected Years of School	1,000	0,870
Mean Years of School	1,000	0,913
Adjusted of Per Capita Expenditure	1,000	1,000

Extraction Method: Principal Component Analysis.

The communality table shows the value of the variable whether it can explain a factor or not. Variables are considered able to explain the factor if the Extraction value is greater than 0.50. Based on the above output, it is known that the extraction value for all variables is greater than 0.50. Thus it can be concluded that all variables can be used to explain factors.

Tabel 3. The Result of Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.360	83.993	83.993	3.360	83.993	83.993
2	.289	7.234	91.227	.289	7.234	91.227
3	.217	5.432	96.659			
4	.134	3.341	100.000			

Extraction Method: Principal Component Analysis.

The Explained Total variance table shows the value of each variable analyzed. In this study, there are four variables meaning there are four variables analyzed. There are two types of analysis to explain variants, Initial Eigen Value and Extraction Sums of Squared Loadings. Initial Eigenvalues indicate the factors formed. Whereas the Extraction Sums of Squared Loadings section shows the number of variations or many factors that can be formed.
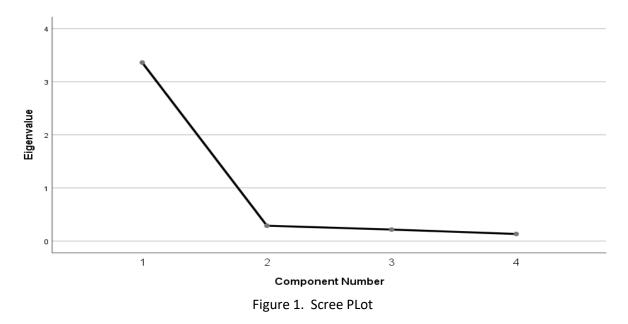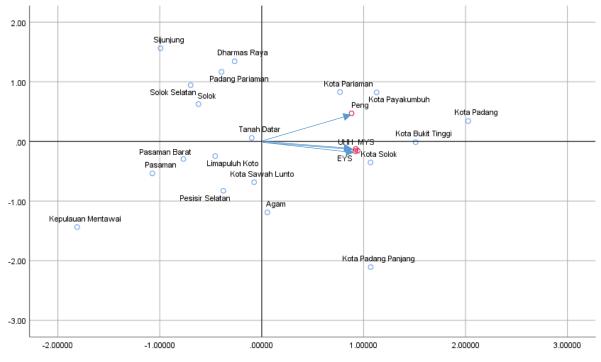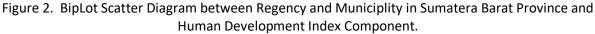

Figure 1. Scree PLot

Figure 1. shows the Scree plot, it shows that the root features of the first and second variables are more than 2. Thus it can be concluded that there are two factors that can be used in this study.

Figure 2. BipLot Scatter Diagram between Regency and Municiplity in Sumatera Barat Province and Human Development Index Component.

T

he Figure 2 ilustrates the biplot scatter diagram, there are several important findings revealed from this study, including:

1. Relative similarities between objects

    The resulting biplot image shows districts / cities in West Sumatra scattered in all quadrants. The groups can be divided as follows:

    a. Pasaman Barat Regency, Pasaman Regency, Lima Puluh Kota District, Sawahlunto Municipaty, Pesisir Selatan Regency, Agam Regency.

    b. Sijunjung Regency, Solok Regency, South Solok Regency, Padang Pariaman Regency, Dharmasraya Regency, Tanah Datar Regency.

    c. Pariaman Municipality and Payakumbuh Municipality.

    d. Bukitingg Municipality i and Solok Municipality.

    e. Meanwhile PadangMjnicipality is separate apart, as well as Padang Panjang Municipality, and Mentawai Regency.

2. Relationships between variables

    Life Expectancy (UHH), School Expectancy Rate (AHS), and Average Length of School (RLS) as well as adjusted expenditure show a positive correlation. Adjusted expenses have the weakest relationship compared to other variables.

3. Variable diversity,

    The most diversity variable is adjusted expenditure per capita. It can be indicated from the vector length, it length is almost 1. Meanwhile the lowest diversity occurs in the life expectancy variable.

4. The relationship between variables and observed objects.

    Figure 2 also presents a general description of the relationship between observations and variables, or in this research, the relationship between regency / municiplity and HDI's variable. This relationship can also be grouped as follows:

    a. Pasaman Barat Regency, Pasaman Regency, Lima Puluh Kota Regency, Sawahlunto Municipality, Pesisir Selatan Regency, and Agam Regency have similarities in terms of excellence in Expected Years of School (EYS), then Life Expectancy (LE), then Mean Years of School (MYS), and finally is adjusted per capita expenditure.

    b. Sijunjung Regency, Solok Regency, Solok Selatan Regency, Padang Pariaman Regency, Dharmasraya Regency, and Tanah Datar Regency have similarities in terms of adjusted per capita expenditures, Life Expectancy (LE), Expected Years of School (EYS), and Mean Years of School (MYS).

    c. Pariaman Municipality and Payakumbuh Municipality are located in the same direction as the adjusted expenditure variable, so that those observation are veri hight in adjusted per capita expenditures, then other variables follow, successively Life Expectancy (LE), Mean Years of School (MYS), and Expected Years of School (EYS).

    d. Solok Municipality and Bukitinggi Municipality are in the same direction as the Mean Years of School (MYS), Life Expectancy (LE), Expected Years of School (EYS), and adjusted per capita expenditures.

    e. There are three regions that are significantly different compared to other regions. Padang Municipality is in area of adjusted per capita expenditures, so that this municipality is in advance for that variable, however Padang Municipality is also very good in Expected Years of School (EYS).

    f. Padang Panjang Municipality is in line with Mean Years of School (MYS), it can be concluded that region is very good in that variabel extremly.

    g. Mentawai Islands Regency is in another zone, in the opposite direction to all variables, as a representation it has the lowest value on all variables.

In the usual studies, the main discussion of HDI is about regional ranking One weakness when focussing on ranking can obscure the potential of the region. Other studies

that have been carried out link HDI with other factors. Research like this clearly has a different purpose, which means to increase the HDI then those factors will be corrected. This paper is the first subject to examine HDI using biplot analysis in Sumatra Barat Province. In this study the supporting dimensions of the HDI are discussed. As stated above, the relative similarity between objects, relationships between variables, the variable diversity and the relationship between variables and observed objects. Understanding this situation bring to a more interesting portrait of the HDI and in the Province of Sumatra Barat.

CONCLUSIONS

The objective of this study is to analyze the strength of the HDI variables in each region in Sumatera Barat Province, using Biplot Analysis. The results show that the HDI variables do not spread in all directions in the biplot diagram, however they fall in the right section of the scatter diagram. Important findings resulted from the study are as follow:

1. The study area is classified into five regional groups that are affected by the proximity value of its constituent variables.
2. Two groups are in an area not close to the variable. The other two groups are in the variable area, while the last group is an extreme region compared to other regional groups. For extreme groups, Kota Padang is very superior in the adjusted per capita expenditure variable and superior in other variables. Kota Padang Panjang is very superior in all variables of education but weak in the economic variable. Meanwhile, the Mentawai Islands Regency is very weak in all variables.
3. The areas which are not in the area of variables are weak against variables, while the area within the variable area is superior in the relevant variable.

This study concludes that biplot is capable to explain the strenghtness of the variable that supports HDI. Thus, it is recomended to use this method in analysis the HDI of a region.

ACKNOWLEDGEMENT

REFERENCES

BPS.(2019). *Indek Pembangunan Manusia*. Jakarta: Badan Pusat Statistik.

Junian, R., Kusnandar, D., & Sulistianingsih, E. (2018). Analisis Indeks Pembangunan Manusia Di Kalimantan Barat Dengan Regresi Panel Dan Biplot. *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster), 07*(03), 177 – 184.

Hatidja, D., & Momuat, L., (2017). Penggunaan Analisis Biplot Untuk Memetakkan Sekolah Dasar Di Kabupaten Minahasa Tenggara Berdasarkan Indikator Stándar Nasional Pendidikan. *D'cartesian:Jurnal Matematika dan Aplikasi*, 6(2),96-108.

Leleury, Z., & Wokanubun, A., (2015). Analisis Biplot pada Pemetaan Karakteristik Kemiskinan di Provinsi Maluku. *Jurnal Ilmu Matematika dan Terapan*,9(1),21-31.

Prasetyoningrum, A., & Sukmawati, S. (2018). Analisis Pengaruh Indeks Pembangunan Manusia (IPM), Pertumbuhan Ekonomi dan Pengangguran Terhadap Kemiskinan di Indonesia. *Equilibrium: Jurnal Ekonomi Syariah*, 6(2), 217-240.

Sholihin, I., Mustafid, & Safitri, D. (2014). Faktor Konfrimatoristragegi Positioning Pasar Modern Indomaret. (Studi Kasus Wilayah Tembalang Kota Semarang). *Jurnal gaussian*, 3(3), 431–440.

Tinungki, G. (2018). AccuracyLevel OfBiplot Analysis Based On The Variance-Covariance Matrix. *American Journal of Engineering Research*, 7(5),264-271.