Conference Paper

Analysis of Human Development Index of 13 Districts in West Kalimantan

Evy Sulistianingsih, Risma Junian, and Setyo Wirarizki

University of Tanjungpura, Kota Pontianak, Kalimantan Barat 78115, Indonesia

Abstract

Human Development Index (HDI) is an indicator that becomes a prominent aspect reflecting the achievement of income, education, and health of community in a region. West Kalimantan, one of the provinces in Indonesia, is ranked 29 out of 34 provinces in terms of HDI. In addition, HDI of the province is the worst one among other five provinces in Kalimantan. For that reason, it is crucial to conduct research on influential factors that affect West Kalimantan's HDI. There are three factors projected to give impact to HDI in this research. Those factors are Labor Force of Participation Rate (LFPR), Population Density (PD), and Poverty Level (PL). Data of the three factors from 13 districts in the province will be analyzed by panel regression and biplot. Panel regression used in this research assumes that there is no time-specific effect, slope coefficients are constant, and the intercepts vary across individuals. According to the result of analysis, it can be summarized that the fixed effect with adjusted determination-coefficient of 0.69 is the best model to analyze this case, where the three factors are statistically significant to HDI of the districts. After getting variables influencing HDI, biplot analysis with alpha 0 was conducted to the data. The latter analysis concluded that there was a strong and positive correlation between PL and LFPR. Moreover, the biplot analysis summarized that Sambas, Bengkayang, Mempawah, Sanggau, Sintang, Kapuas Hulu, Sekadau, and Melawi have a bigger number of LFPR rather than the other districts.

Keywords: longitudinal, descriptive, singular-value-decomposition

Corresponding Author:
Evy Sulistianingsih
evysulistianingsih@math.untan.ac.id

Received: 29 August 2018
Accepted: 18 September 2018
Published: 11 November 2018

Publishing services provided by Knowledge E

Selection and Peer-review under the responsibility of the ICOI-2018 Conference Committee.

1. Introduction

Human Development Index (HDI) is an indicator utilized to measure quality of human life in a region. HDI becomes a prominent aspect that can reflect achievement of income, education, and health of community.

According to BPS-Statistics Indonesia (2017), West Kalimantan, South Kalimantan, and also Central Kalimantan, three of the provinces in Indonesia, were consecutively ranked 28, 20, and 19 out of 35 provinces in terms of HDI. Meanwhile, East Kalimantan was ranked the third one among 35 provinces in Indonesia in terms of HDI. Based

on that fact, there was a big gap of HDI condition among all provinces situated in Kalimantan. For that reason, conducting a research about some factors considered having influence to HDI of West Kalimantan becomes crucial.

Labor Force of Participation Rate, Population Density, and Poverty Level will be used in this research as some factors that are predicted can give influence to HDI. Data of the factors from chosen districts in West Kalimantan will be analyzed by panel regression and Biplot.

Panel regression used in this research is assumed that there is no time specific effect, slope coefficients are constant, and the intercepts varies over individuals. The best panel regression model will be obtained after checking the results of Chow, Hausman, and Lagrange Multiplier Test. Those tests are utilized to decide an appropriate model among common effect, fixed effect, and random effect. Furthermore, this research would give important implications for the researchers and other researchers concerned with panel regression, biplot and also HDI analyses.

2. Literature Review

There were some researchers such as Heriyanto (2012) and Ayunanda and Zain (2013) who analyzed if some factors, could influence HDI of some districts. Heriyanto (2013) examined Gross Domestic Product (GDP), Social Development Budget, Poverty Level, and Facility of Health and Education that were predicted to have influenced Human Development Index (HDI) in West Kalimantan between 2006 and 2010 by panel regression analyses. The research summarized that GDP, Social Development Budget, and Facility of Health and Education significantly gave positive impact to HDI while Poverty Level significantly contributed negative impact to HDI of West Kalimantan.

Besides Heriyanto (2012), Ayunanda and Zain (2013)also utilized panel regression to analyze some influential factors to HDI of East Java from 2004 until 2011. Based on the research, it could be summed up that there were seven factors that were statistically significant in giving impact to East Java's HDI. Two of them were Labor Force of Participation Rate and Population Density. Furthermore, this research also concluded that Fixed Effect was the best model to analyze this case with adjusted-determination coefficient of the model at 0,96.

In addition, many previous researchers have been executed to analyze HDI in many regions by mixture of two methods; panel regression and biplot analysis. Some of the researchers who used this combination were Rustariyuni (2014), Destilunna and

Zain (2015), Nurhasanah, Salwa, and Amelia (2016) and Heriyanto and Kinansi (2013). Reviews of their research will be explained as follows.

Firstly, Rustariyuni (2014) conducted a study to analyze HDI using Pooled Least Square. The research focused to analyze HDI of Bali from 2004 until 2012.Some analyzed factors in this research were gini *ratio*, non-food expenditure percapita, local budget, and economic growth rate. According to the research, it can be concluded that gini ratio, non-food expenditure percapita, local budget, and economic growth rate significantly donated positive impact to IPM.

Destilunna and Zain (2015) conducted research that focused on getting a model that could describe some influential factors to HDI in East Java. The research used panel data from 2008 until 2012. This study not only analyzed some major factors that influenced the HDI, but also resulted in a biplot analysis that constructed a map of variables affecting the HDI. According to the research, Fixed Effect Method (FEM) with a determination coefficient at 98.88% was the best model to analyze this case. Furthermore, the research also showed that the ratio of student and teacher (RST) for Junior High School and Islamic Senior High School, Rate of School Participation (RSP) for Junior High School Student, RSP for Senior High School Student, Growth Domestic Product (GDP) of Small Micro Medium Enterprise (SMME), and also Open Unemployment Rate were statistically significant to influence HDI of East Java. In addition, the result of biplot map summarized that an increase of SMME's GDP of the lower middle class could be created by an increase of RST for Junior High School Student and Islamic Senior High School.

Panel regression and biplot analysis were used also by Nurhasanah, Salwa, and Amelia (2016) in order to make a description of touristic characteristic of districts in Aceh Province. The description derived from some variables of touristic product and to determine a model of tourist number for each district in Aceh Province between 2008 and 2013. According to biplot result, districts stood on the first and second rank in terms of the total amount of tourism were Sabang and Banda Aceh. Furthermore, according to the result of panel regression analyses, it could be concluded that FEM was the best model rather than Common Effect Method and Random Effect Method. The FEM showed that number of tourists in each districts was influenced by accommodation number, restaurant number, and tourism place number.

In addition, biplot analysis also was used by Heriyanto and Kinansi (2013) to get information about the kind of diseases that became the ultimate diseases in both twelve sub-districts in Batam and twenty one districts in Nusa Tenggara Timur Province in 2009. The research summarized that sub-district of SeiBeduk; Lubuk Baja and Batam

City were sub-districts which had a higher level of Measles rather than the other sub-districts. The research also showed that two sub-districts namely Batu Aji and Nongsa were the places where the biggest number of cases of Sexually Transmitted Disease (STD) occurred. Meanwhile, biplot analyses for Nusa Tenggara Timur Province concluded that Kupang and Sumba Timur were the districts that needed more intense concern from the government because of so many lethal diseases assembled at the two sub-districts namely STD, Tuberculosis, Measles, Malaria, Diarrhea, and Dengue Fever.

There were some evidence that panel regression and biplot analysis could be used to analyze those aforementioned cases. Furthermore, there were facts that Labor Force of Participation Rate, Population Density, and Poverty Level were statistically significant to influence HDI in other districts in Indonesia. Therefore, this article intends to analyze the relationship between the three factors and 13 district's HDI in West Kalimantan: Sanggau, Sintang, Sambas, Ketapang, Bengkayang, Landak, Singkawang, Kapuas Hulu, Melawi, Sekadau, Pontianak, and Kayong Utara by using panel regression and biplot. This research will analyze district-level data of those factors from 2008 to 2015 derived from BPS of West Kalimantan (2008–2016).

3. Methods

This research used both panel regression and biplot to analyze the estimated factors influencing HDI of the districts. Those methods were explained briefly as follows.

Hsiao [2014] claimed that panel regression model could be more accurate to estimate parameters of a regression model because there are a large number of data points in panel regression model. Baltagi [2005] expressed a panel regression model (PRM) as follows

$$z_{it} = \alpha_{it} + \beta_{it} x_{it} + \varepsilon_{it} \;\; i = 1, 2, \ldots, \; N \;\; t = 1, 2, \ldots, \; T, \tag{1}$$

where z_{it} is dependent variable, x_{it} is independent variable, and ε_{it} is the error term, uncorrelated with x_{it}, with mean zero and constant variance σ_{ε}^2. Furthermore, α_{it} is a scalar and β_{it} is regression coefficient, *slope*. On equation (1), the subscript i denotes the cross section dimension whilst $t$symbolises the time series dimension. Baltagi [2005] also stated that PRM concentrates to control the impact heterogeneity that cannot be unobserved. This way was conducted to get valid inference on β_{it}. For example,

unobserved heterogeneity in a linear regression model is assumed as an individual specific and time invariant. So that (1) can be expressed as

$$z_{it} = \alpha_i^* + \beta_i x_{it} + \varepsilon_{it} \;\; i = 1, 2, \ldots, \; N \;\; t = 1, 2, \ldots, \; T. \tag{2}$$

The parameters α_i^* and β_i for different cross sectional units can be different, even though both of them stay constant over time. Then, sampling distribution can seriously mislead the least square regression of z_{it} on x_{it} when all of NT observations are used to estimate (3) as follows [Hsiao, 2014]

$$z_{it} = \alpha^* + \beta x_{it} + \varepsilon_{it} \;\; i = 1, 2, \ldots, \; N \;\; t = 1, 2, \ldots, \; T. \tag{3}$$

Sulistianingsih et al. [2017] argued that there are three models that can be chosen to obtain estimators in panel regression, namely Common Effect Model (CEM), Fixed Effect Model (FEM), and Random Effect Model (REM). Those three models can be formulated as written at [Hsiao, 2014].

Moreover, panel regression model involves some statistical tests to analyze independent variables and dependent variables. The tests are Chow Test, Hausman Test and also Lagrange Multiplier Test. The function of the three statistical tests and the methodology to conduct panel regression analysis was mentioned at Ayunanda and Zain [2013].

Some steps included in analyzing panel data were written briefly as follows (Ayunanda and Zain, 2013):

1. Analyzed the characteristics for independent and dependent variables
2. Checked multicollinearity among independent variables
3. Conducted Chow Test, Hausman Test, and Lagrange Multiplier Test
4. Chose the best model among CEM, REM and FEM according to result of Step 3
5. Checked variables which are statistically significant at the chosen model at Step 4
6. Conducted Heteroscedasticity Test, Autocorrelation Test, and Normality Test

After getting some variables that could influence 13 districts' HDI in West Kalimantan, biplot analysis was conducted. Biplot analysis itself was a statistical descriptive technique that could form a representative graph presenting *n* objects of observation and *p* variables that were relatively corrected of its average on a two-dimension graph (Jolliffe, 2002). In addition, according to Mattjik and Sumertajaya [2011], there were

four important functions that could be seen from biplot namely proximity among observed objects, variety of variables, correlation among variables, and score of variable related to an object.

Methodology of biplot analysis based on SVD in this research were studied extensively from Leon [2001], Mattjik and Sumertajaya [2011], Joliffe [2010], Gabriel (1971), and Heriyanto and Kinansi (2013). Biplot analysis was executed by several steps as follows:

1. Singular Value Decomposition (SVD)
   (a) Determined average value of each independent variable from 2008 until 2015 for each observed district in West Kalimantan.
   (b) Arranged observation in a matrice X**X** which contains of *n* objects and *p* variables.
   (c) Transformed matrice X**X** becomes matrix **X*** by standardization of the data. For simplification purpose, **X** was reused to substitute **X*.**
   (d) Calculated matrice $X^{'}X$ and its Eigen value and Eigen vector to get matrices **U**, **L**, and **A.**
2. Biplot Analyses
   (a) Constructed a row matrice **G** = **U** $\mathbf{L}^{\alpha}$ and column matrice $H^{'}=L^{1-\propto}A^{'}$. **G** was a coordinate point of 13 districts and $H^{'}H'$ was coordinate point of three independent variables.
   (b) Took the first-two columns of Gand the first-two rows of $H^{'}H'$to become matrices $\mathbf{G}_2$ and $\mathbf{H}_2$, where each row of $\mathbf{G}_2$ was a coordinate point (*x*,*y*) for each district. Meanwhile each row of $\mathbf{H}_2$ was a coordinate point (*x*,*y*) for each independent variable. Then the latter matrices was used to develop biplot graph.
   (c) Measured goodness of fit of biplot.
   (d) Interpreted the biplot.

4. Results and Discussion

In this section, panel regression was used to examine district-level data from 2008 to 2015 of three factors namely Poverty Level (x_1), Population Density (x_2), and Labor Force of Participation Rate (x_3), which were considered giving influence to HDI of

the 13 districts namely Sanggau, Sintang, Sambas, Ketapang, Bengkayang, Landak, Singkawang, Kapuas Hulu, Melawi, Sekadau, Pontianak, and Kayong Utara. Before analyzing the data by panel regression, multicollinearity test was conducted previously to check whether the multicollinearity had happened or not. According to the result of multicollinearity test, VIF value for each variable was smaller than 10. Hence, there was no multicollinearity between the independent variables. After ensuring that there was no multicollinearity among the independent variables, Chow Test was done to evaluate which one was better between CEM and FEM to be used in modelling the panel data.

After Chow Test, Hausman Test was conducted to examine whether there was a random effect or not in the model. According to Hausman Test's result, FEM was a better model than REM. It was shown by chi-square-statistics, 23.676, which was bigger than chi-square table, with degrees of freedom 3 and $\alpha = 5\%$, 7.815.

Next, Lagrange Multiplier Test was executed to check heteroscedasticity in the model. Then, based on Lagrange Multiplier Test, it could be deduced that there was no heteroscedasticity in FEM.

According to the three test results, it could be concluded that FEM was the best model than CEM and FEM. Furthermore, based on the results of Heteroscedasticity Test, Autocorrelation Test, and Normality Test, it could be summarized that the FEM fulfilled heteroscedasticity, autocorrelation, and normality assumption. So that, this model could be utilized to describe the relationship between dependent and independent variables utilized in the model. Moreover, FEM could distinguish effect of every district, so estimation of HDI for 13 districts in West Kalimantan could be written as follows at Table 2:

Table 2 shows that there were various intercepts among 13 districts in West Kalimantan.

After getting the three factors which were statistically significant to HDI. data from the three factors was standardized and then arranged in a matrice X, as follows:

Table 1: FEM for 10 districts in West Kalimantan.

District	Estimation of HDI
Sambas	$84.11969 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Bengkayang	$86.47784 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Landak	$91.65719 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Mempawah	$83.1474 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Sanggau	$82.31139 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Ketapang	$88.96579 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Sintang	$87.7688 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Kapuas Hulu	$89.14781 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Sekadau	$84.05377 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Melawi	$91.59305 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Kayong Utara	$84.51093 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Pontianak	$102.6946 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$
Singkawang	$84.0846 - 0.900871x_1 - 0.003476x_2 - 0.107718x_3$

$$X = \begin{pmatrix} 0.198 & -0.272 & 0.718 \\ -0.494 & -0.298 & 0.293 \\ 1.651 & -0.303 & -0.025 \\ -1.006 & -0.198 & -0.953 \\ -1.388 & -0.304 & 0.444 \\ 1.160 & -0.317 & -0.321 \\ 0.296 & -0.314 & 1.053 \\ 0.417 & -0.321 & 0.790 \\ -0.827 & -0.303 & 0.822 \\ 1.243 & -0.315 & 0.993 \\ 0.649 & -0.311 & -0.183 \\ -0.933 & 3.319 & -2.107 \\ -0.965 & -0.063 & -1.524 \end{pmatrix}$$

Then, matrices U, L, and A were formed by Singular Value Decomposition (SVD). Those matrices were written as follows:

$$L = \begin{pmatrix} 4.809 & 0 & 0 \\ 0 & 3.003 & 0 \\ 0 & 0 & 1.963 \end{pmatrix}; A = \begin{pmatrix} -0.463 & -0.883 & -0.079 \\ 0.619 & -0.385 & 0.685 \\ -0.635 & 0.268 & 0.724 \end{pmatrix};$$

$$U = \begin{pmatrix} -0.149 & 0.041 & 0.162 \\ -0.030 & 0.21 & 0.024 \\ -0.194 & -0.449 & -0.181 \\ 0.197 & 0.236 & -0.380 \\ 0.036 & 0.486 & 0.113 \\ -0.111 & -0.329 & -0.276 \\ -0.208 & 0.047 & 0.267 \\ -0.186 & -0.011 & 0.163 \\ -0.068 & 0.355 & 0.231 \\ -0.291 & -0.236 & 0.207 \\ -0.078 & -0.167 & -0.202 \\ 0.795 & -0.339 & 0.418 \\ -0.286 & 0.156 & -0.546 \end{pmatrix}$$

Next, matrices $\mathbf{G}_2$ and $\mathbf{H}_2$ were formed with$\alpha = 0$,

$$G_2 = \begin{pmatrix} -0.149 & 0.041 \\ -0.030 & 0.21 \\ -0.194 & -0.449 \\ 0.197 & 0.236 \\ 0.036 & 0.486 \\ -0.111 & -0.329 \\ -0.208 & 0.047 \\ -0.186 & -0.011 \\ -0.068 & 0.355 \\ -0.291 & -0.236 \\ -0.078 & -0.167 \\ 0.795 & -0.339 \\ -0.286 & 0.156 \end{pmatrix}; H_2 = \begin{pmatrix} -2.225 & -2.652 \\ 2.975 & -1.156 \\ -3.054 & 0.805 \end{pmatrix}$$

Every row of matrice $\mathbf{G}_2$ was the coordinate point (*x*,*y*) of each district whilst every row of matrixe$\mathbf{H}_2$ was the coordinate point (*x*,*y*) of each independent variable. Then, both matrices $\mathbf{G}_2$ and $\mathbf{H}_2$ were utilized to construct Figure 1 and to interpret the result of biplot analysis where x[,1] was the coordinate point on *x* for the first column of $\mathbf{G}_2$ and $\mathbf{H}_2$ while x[,2] was the coordinate point of *y* for the second column of $\mathbf{G}_2$ and $\mathbf{H}_2$. Moreover 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12 represented consecutively for Sambas, Bengkayang, Landak, Mempawah, Sanggau, Ketapang, Sintang, Kapuas Hulu, Sekadau, Melawi, Kayong Utara, Pontianak, and Singkawang.

According to the analysis, it showed that the goodness of fit of the biplot for this case was 0.8930. This number described that there was approximately 89.30% variations of independent variables that could be explained from the first two columns of matrice **G** and **H.** In other words, biplot analysis in this research gave good presentation in order to describe information related to variables influencing HDI of the districts.

Based on biplot analysis, it could be calculated that the longest vector among the three independent variables was vector of Poverty Level, 3.460. This number showed

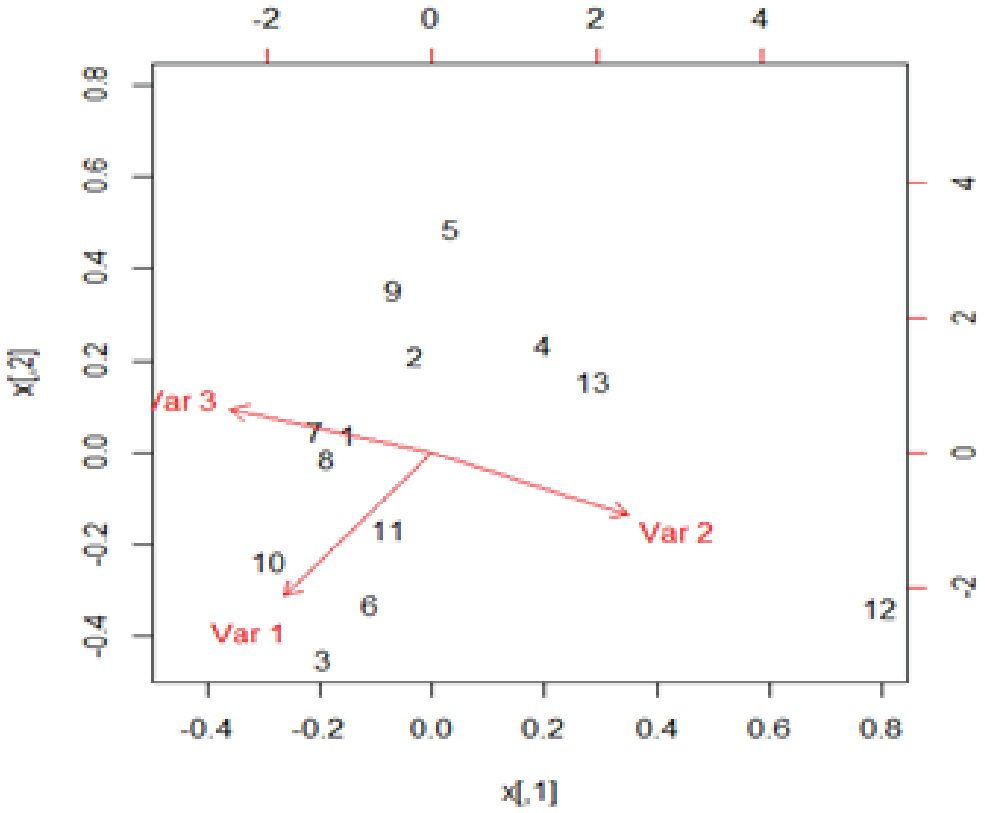

Figure 1: Biplot result of factors influencing HDI of West Kalimantan.

that poverty level did not spread equally in each district of West Kalimantan. Meanwhile, the shortest vector among the three variable vectors is vector of Labor Force of Participation Rate, 3.158. This latter result described that level of Labor Force of Participation Rate spread equally in each observed district.

Furthermore, from biplot analysis, it could be also calculated and concluded that Landak, Ketapang, Kayong Utara, Melawi, and Kapuas Hulu had a bigger number of the poor rather than the other districts because there were narrow angles which were formed between the latter variable and each district.

5. Conclusion

According to the panel regression analysis, Poverty Level, Population Density, and Labor Force of Participation Rate (x_3) are statistically significant to influence Human Development Index of the 13 districts in West Kalimantan. The biplot analysis results

summarize that the goodness of fit for the biplot analysis of this case is 0.8930. Therefore, biplot analysis in this research gives good presentation in order to describe information related to variables influencing HDI of the districts. The latter analysis summarized that poverty level did not spread equally in each district of West Kalimantan. Meanwhile, level of Labor Force of Participation Rate spread equally in each observed district.

Acknowledgement

The authors would like to thank the Faculty of Mathematics and Natural Sciences Faculty for providing the financial support for this research work and the referees for the comments that greatly improved the manuscript.

References

[1] BPS-Statistics Indonesia. (2017). *Human Development Index 2016*. BPS-Stat.

[2] Heriyanto, D. (2012). Analisis Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia (IPM) Kabupaten/Kota di Provinsi Kalimantan Barat Tahun 2006-2010. *Jurnal Ekonomi Daerah*, vol. 1, no. 1, pp. 1–18.

[3] Ayunanda, M. and Zain, I. (2013). Analisis Statistika Faktor yang Mempengaruhi Indeks Pembangunan Manusia di Kabupaten/Kota Provinsi Jawa Timur dengan Menggunakan Regresi Panel. *Jurnal Sains dan Seni POMITS*, vol. 2, no. 2, pp. 237–242.

[4] Rustariyuni, S. D. (2014). Pengaruh Gini Ratio, Pengeluaran Non Makanan per Kapita, Belanja Daerah dan Laju Pertumbuhan Ekonomi Pada Indeks Pembangunan Manusia Kabupaten/Kota di Provinsi Bali Periode 2004-2012. *Jurnal Piramida*, vol. X, no. 1, pp. 45–55.

[5] Destilunna, F. G. and Zain, I., (2015). Pengaruh dan Pemetaan Pendidikan, Kesehatan, serta UMKM terhadap Indeks Pembangunan Manusia di Jawa Timur Menggunakan Regresi Panel dan Biplot. *Jurnal Sains dan Seni ITS*, vol. 4, no. 2, pp. 292–298.

[6] Nurhasanah, S. N. and Nelva, A. (2016). Penentuan Karakteristik Pariwisata dan Model Jumlah Wisatawan untuk Kabupaten/Kota di Provinsi Aceh. *Jurnal Natural*, vol. 16, no. 1, pp. 43–50.

[7] Heriyanto, B. and Kinansi, R. R. (2013). Analisis Biplot pada Data kasus Penyakit di Beberapa Daerah di Indonesia Tahun 2009. *Bulletin Penelitian Kesehatan*, vol. 41, no. 2, pp. 120–130.

[8] BPS of West Kalimantan. (2014–2016). *Human Development Index*. Pontianak: BPS of West Kalimantan.

[9] Baltagi, B. H. (2005). The one-way error component regression model, in *Econometric Analysis of Panel Data* (third edition). England: The Atrium, John Wiley and Sons, Ltd.

[10] Hsiao, C. (2014). Simple Regression with Variable Intercepts, in *Analysis of Panel Data* (third edition). USA: Cambridge University Press.

[11] Sulistianingsih, et al. (2017). Analysis of palm oil production, export, and government consumption to gross domestic product of five districts in West Kalimantan by panel regression. *Journal of Physics: Conference Series*, vol. 824, no. 1.

[12] Jollife, I. T. (2010). *Principal Component Analysis* (second edition). New York, NY: Springer.

[13] Mattjik, A. A. and Sumertajaya, I. M. (2011). *Sidik Peubah Ganda*. Bogor: IPB Press.

[14] Chalid, N. (2014). Pengaruh Tingkat Kemiskinan, Tingkat Pengangguran, Upah Minimum Kabupaten/Kota dan Laju Pertumbuhan Ekonomi terhadap Indeks Pembangunan Manusia di Riau. *Jurnal Ekonomi Universitas Riau*, vol. 22, no. 2, pp. 1–12.

[15] Gabriel, K. R. (1971). The Biplot graphics display of matrices with applications for principal component analysis. *Biometrika*, vol. 58, no. 3, pp. 453–467.