Implementation of Biplot Analysis for Mapping Elementary and Junior High Schools in West Sumatra Based on National Examination Results 2016

N Amalita*, D Fitria, V Distian

Mathematics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Padang, Jl. Prof. Hamka, Padang, Sumatera Barat, Indonesia 25131

*nongaditya@gmail.com

Abstract. National examination is an assessment of learning outcomes that aims to assess the achievement of graduate competence nationally. The result of the national examination is used as a mapping of educational issues in order to arrange the national education policy. Therefore the results of National Examination are used, also, as a reference for the admission of new students to continue their education to a higher level. The results of National Examination in West Sumatra in 2016 decreased from the previous year, both elementary schools (SD) and Junior High School level (SMP). This paper aims to determine the characteristics of the National Examination results in each regency/ city in West Sumatra for elementary and junior levels by using Bi-plot analysis. The result of Bi-plot Analysis provides the information that the results of the National Examination of Regency/ City in West Sumatra Province are quite diverse. At Junior High School level there are 9 of Regencies/ Cities which have similar characteristics. English subjects are the greatest diversity among all of subjects. The calculation results of the correlation of each variable in junior high school level are positively correlated. The variables with positive correlation are mathematics that correlates with English. Based on the mark of National Examination for elementary school level in West Sumatra, there are 8 Regencies/ Cities have similar characteristics. The correlations of each variable at the elementary level are positively correlated. The variables that have positive correlation are Sciences (IPA) with Language.

1. Introduction

One of the objectives of the National Exam (UN) is to assess the achievement of national graduate competence on certain subjects in science and technology groups. The National Exam is a test conducted in the last year of each educational unit. Until now the UN is still being implemented although there are still pros and cons and society. The meaning of the National Examination is the evaluation system of basic and intermediate education standards nationally and the quality equation of education between regions conducted by the Center for Educational Assessment, Depdiknas in Indonesia based on the Law of the Republic of Indonesia number 20 of 2003.
Benefit UN in BSNP 2016, is

- Mapping the quality of educational programs and / or educational units
- Selection consideration of the next level of education; and

- The basis of guidance and provision of assistance to educational units for equity and improvement of education quality.

Although the value of the UN is not a graduation of a student, the UN is still a guideline for continuing education to a higher level, so for the students the value of the UN is very important and should strive to improve the test results in order to get a favorite school. While for the school is a pride and as one indicator as a quality school and excel. Not all subjects conducted National Examination, for Elementary School level there are three subjects namely Mathematics, Bahasa and Science. For Junior High School the subjects are Indonesian, Science, Mathematics, and English

Furthermore, the average value of junior high school and elementary level in the regency / city in West Sumatra can be seen in the following table:

Table 1: Average value of junior high school and elementary level in the regency / city in West Sumatra

No	Name of District / City	Average National Exam	
		SD	SMP
1	Padang City	75,81	66,96
2	Bukittinggi City	83,67	64,51
3	Padang Panjang City	74,38	60,12
4	Sawahlunto City	77,55	55,98
5	Solok City	55,35	54,31
6	Payakumbuh City	75,65	59,48
7	Pariaman City	75,04	55,67
8	Agam District	79,55	52,44
9	Pasaman District	72,34	49,42
10	Lima Puluh Kota District	77,92	52,91
11	Solok District	71,82	50,03
12	Padang Pariaman District	68,87	48,51
13	Pesisir Selatan District	73,69	44,03
14	Tanah Datar District	75,58	56,01
15	Sijunjung District	71,64	49,99
16	Kepulauan Mentawai District	60,08	53,76
17	Pasaman Barat District	67,63	47,58
18	Solok Selatan District	78,89	49,18
19	Dharmasraya District	75,54	50,31

In the table above shows that the average value of the UN for the highest junior high school is the city of Padang. While for the lowest value in the District of West Pasaman and Pesisir Selatan district. The highest average grade of primary school in Bukittinggi and the lowest in Solok city. It means that the value of UN both for primary and junior high school level is in big city. For that reason, it is necessary to mapping based on the result of National Examination. These results can be used to evaluate and improve the results of the National Exam the following year. Mapping used Biplot Analysis. Biplot analysis is one multivariate analysis technique that can be used to solve the problem. According to Mattjik (2011: 247), as for information that can be obtained from the analysis of biplot is to see the proximity between objects, diversity variables, relationships between variables, and the

value of the variable on an object. The object of this research is 19 districts / cities in West Sumatera Province. Meanwhile, the variables in the study are subjects tested in the National Exam.

2. Method

Biplot Analysis will be employed in this work. Biplot is a graphical technique that presents together information about the object of observation and variables in a plot (Mattjik & Sumertajaya, 2011: 246). If the translation of the X data matrix is sized (n × p), then the **X** matrix can be decomposed into:

$$\mathbf{X}_{(n \times p)} = \mathbf{G}_{(n \times r)} \mathbf{H}^t_{(r \times p)}$$

Biplot analysis, data analysis will be performed on the correlated data matrix against the mean. The corrected matrix of the mean is denoted by $\widetilde{\mathbf{X}}$ and can be written The corrected matrix of the mean is

denoted by $\widetilde{\mathbf{X}}$ and can be written is $\widetilde{\mathbf{X}} = \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \cdots & \tilde{x}_{1p} \\ \tilde{x}_{21} & \tilde{x}_{22} & \cdots & \tilde{x}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{n1} & \tilde{x}_{n2} & \cdots & \tilde{x}_{np} \end{bmatrix}$

If the matrix $\widetilde{\mathbf{X}}$ is a standardized data matrix, then based on the decomposition of the singular value of X matrix can be described as:

$$\widetilde{\mathbf{X}}_{(n \times p)} = \mathbf{U}_{(n \times r)} \mathbf{L}_{(r \times r)} \mathbf{A}^t_{(r \times p)} = \mathbf{G}_{(n \times r)} \mathbf{H}^t_{(r \times p)}$$
$$\mathbf{G} = \mathbf{U} \text{ dan } \mathbf{H} = \mathbf{AL}^t$$

With the decomponent of the singular value of X data matrix, we get the following matrix U:
$\mathbf{U} = \widetilde{\mathbf{X}}(\mathbf{A}^t)^{-1}\mathbf{L}^{-1}$, matriks **A** is ortogonal ortogonal with $\mathbf{A}^{-1} = \mathbf{A}^t$ so $\mathbf{U} = \widetilde{\mathbf{X}}\mathbf{AL}^{-1}$ so that G and H is

$$G = \begin{bmatrix} \lambda_1^{-1/2}(x_1 - \overline{X})^t \boldsymbol{a_1} & \lambda_2^{-1/2}(x_1 - \overline{X})^t \boldsymbol{a_2} & \dots & \lambda_r^{-1/2}(x_1 - \overline{X})^t \boldsymbol{a_r} \\ \lambda_1^{-1/2}(x_2 - \overline{X})^t \boldsymbol{a_1} & \lambda_2^{-1/2}(x_2 - \overline{X})^t \boldsymbol{a_2} & \dots & \lambda_r^{-1/2}(x_2 - \overline{X})^t \boldsymbol{a_r} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{-1/2}(x_n - \overline{X})^t \boldsymbol{a_1} & \lambda_2^{-1/2}(x_n - \overline{X})^t \boldsymbol{a_2} & \dots & \lambda_r^{-1/2}(x_n - \overline{X})^t \boldsymbol{a_r} \end{bmatrix}$$

$$H = \begin{bmatrix} a_{11}\lambda_1^{-1/2} & a_{12}\lambda_2^{-1/2} & \cdots & a_{1r}\lambda_p^{-1/2} \\ a_{21}\lambda_1^{-1/2} & a_{22}\lambda_2^{-1/2} & \cdots & a_{2r}\lambda_p^{-1/2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}\lambda_1^{-1/2} & a_{p2}\lambda_2^{-1/2} & \cdots & a_{pr}\lambda_p^{-1/2} \end{bmatrix}$$

Since the matrix $\widetilde{\mathbf{X}}$ is represented in the field, the matrix to be used as the best approach is the matrix $G^{(2)}$ dan $H^{(2)}$

$$\mathbf{G}^{(2)} = \begin{bmatrix} g_{11} & g_{12} \\ \vdots & \vdots \\ g_{i1} & g_{i2} \\ \vdots & \vdots \\ g_{n1} & g_{n2} \end{bmatrix} \text{ dan } \mathbf{H}^{(2)} = \begin{bmatrix} h_{11} & h_{12} \\ \vdots & \vdots \\ h_{i1} & h_{i2} \\ \vdots & \vdots \\ h_{p1} & h_{p2} \end{bmatrix}$$

The matrices $G^{(2)}$ dan $H^{(2)}$ and are the first two elements of the vector g_i and the vector h_j. The vectors g_i and h_j will be plotted in one plot, this is called the biplot, where g_i will be depicted as a point and h_i will be depicted as the line from the origin (0,0) to the point $\left(a_{j1}\lambda_1^{-1/2}, a_{j2}\lambda_2^{-1/2}\right)$ for $j = 1,2, \dots, p$.

Matjik (2011) gives information and interpretation of biplot analysis are as follows.

- Correlation between each variables described as a directed arrow. Two variables having positive correlation if angle between both variables are acute, and if its angle is obtuse, it means that they have negative correlation and also if they differ in right angle explain that no correlation between both variables.

- Variance of variables sketch as length of vector. Longer the vector, bigger the variances among variables and vice versa shorter the vector, smaller the variances.
- Neighborhood among object is sketched as objects having the same characteristics. Nearest the object, having alike characteristic.
- Variable's value on object detected from the relative position amongst variable and object. Object in a side with variables having the value upper average. If it have opposite direction, the value under average.

3. Result and Discusion

Here are the results of the analysis at junior and primary school levels:

3.1 Junior High School Education Level

Average National Examination of Regency / City level of junior high school education in West Sumatera Province in 2016 is formed into a matrix of data that is **X** matrix that size 19×4. Where, each row in **X** matrix district./city and each column describe the average National Exam per subject

$$
\mathbf{X} = \begin{bmatrix}
75,03 & 69,56 & 57,24 & 65,99 \\
79,81 & 60,97 & 55,05 & 62,21 \\
78,36 & 56,21 & 48,80 & 57,10 \\
76,42 & 52,40 & 42,26 & 52,83 \\
75,64 & 50,40 & 41,40 & 49,78 \\
78,40 & 55,12 & 47,82 & 56,57 \\
72,88 & 56,32 & 45,02 & 48,44 \\
72,47 & 47,26 & 40,16 & 49,87 \\
70,72 & 43,68 & 36,54 & 46,75 \\
74,52 & 45,45 & 39,66 & 51,99 \\
71,66 & 43,48 & 37,20 & 47,38 \\
68,40 & 44,53 & 35,23 & 45,89 \\
63,97 & 39,53 & 32,11 & 40,5 \\
75,61 & 50,42 & 44,48 & 53,51 \\
72,38 & 43,10 & 37,03 & 47,43 \\
64,46 & 55,13 & 46,38 & 49,07 \\
68,80 & 41,42 & 35,01 & 45,09 \\
70,38 & 43,16 & 35,96 & 47,23 \\
71,53 & 44,43 & 36,95 & 48,31
\end{bmatrix}
$$

So matriks **X** is processed so that matriks G^2 dan H^2 is :

$$
G^2 = \begin{bmatrix}
-0,56 & -0,29 \\
-0,42 & 0,15 \\
-0,24 & 0,17 \\
-0,08 & 0,16 \\
-0,01 & 0,12 \\
-0,21 & 0,21 \\
-0,09 & -0,26 \\
0,06 & 0,05 \\
0,18 & 0,03 \\
0,05 & 0,26 \\
0,16 & 0,08 \\
0,20 & -0,12 \\
0,37 & -0,29 \\
-0,08 & 0,16 \\
0,16 & 0,15 \\
-0,06 & -0,69 \\
0,25 & -0,02 \\
0,19 & 0,05 \\
0,14 & 0,08
\end{bmatrix}
\text{ and } H^2 = \begin{bmatrix}
-13,06 & 12,79 \\
-32,04 & -6,63 \\
-29,01 & -2,46 \\
-25,10 & 4,64
\end{bmatrix}
$$

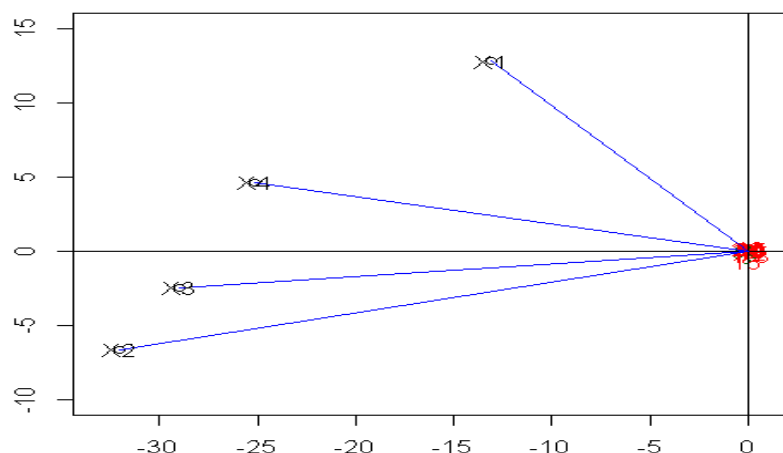Based on G^2 dan H^2 obtain figure biplot is

Figure 1: Biplot Regency / City based on the results of the National Examination in 2016 for Junior High schools level

Based on the above biplot can be taken information about the variance of variables. The English subject variable vector (X2) has a variable vector that is longer than the other variable vectors. The longer the vector variable, the greater the level of variability of the variable. This shows the results of the National Examination of students on English subjects in the Regency / City in West Sumatra Province varied. This means that there are some areas that have a relatively higher average National Examination and some areas have a relatively lower National Examination average.

Based on the biplot above, it can be seen that the vector of Mathematics variable (X3) forms a small angle with the vector of English variable (X2) and also the variable vector of IPA (X4). That is, students who have relatively high Mathematics National Exam score also tend to have relatively high value in English and Science subjects. Meanwhile, the English variable vector (X2) forms a considerable angle with the Bahasa variable vector (X1). This shows that students who have relatively high Bahasa National Examination score, do not always get a relatively high score on English subjects. If the enlarged Biplot graph shows that there are 9 groups :

Table 2: District / City Group based on the similarity of National Examination
Results of 2016 Junior High Schools

Grup	District/City
1	Padang City
2	Bukittinggi City
3	Padang Panjang City
	Payakumbuh City
4	Pariaman City
5	Kepulauan Mentawai District
	Sawahlunto City
6	Tanah Datar District
	Solok City
	Lima Puluh Kota District
	Pasaman District
	Solok District
7	Agam District
	Sijunjung District
	Solok Selatan District
	Dharmasraya District
8	Pasaman Barat District
	Padang Pariaman District
9	Pesisir Selatan District

3.2 Elementary school Level

Average National Examination of Regency / City level of elemetary school education in West Sumatera Province in 2016 is formed into a matrix of data that is $\mathbf{Y}$ matrix that size 19×3. Where, each row in $\mathbf{Y}$ matrix district./city and each column describe the average National Exam per subject.

$$\mathbf{Y} = \begin{bmatrix} 78,96 & 70,95 & 77,53 \\ 84,43 & 81,16 & 85,42 \\ 78,62 & 68,55 & 75,97 \\ 79,65 & 72,76 & 80,25 \\ 79,03 & 69,76 & 77,75 \\ 79,31 & 68,62 & 79,01 \\ 78,38 & 69,44 & 77,30 \\ 80,45 & 76,16 & 82,03 \\ 75,56 & 67,72 & 73,75 \\ 80,30 & 72,80 & 80,66 \\ 75,93 & 66,12 & 73,41 \\ 73,74 & 63,38 & 69,50 \\ 76,69 & 70,19 & 74,18 \\ 79,87 & 69,71 & 77,17 \\ 76,61 & 64,34 & 73,98 \\ 63,10 & 57,46 & 59,68 \\ 73,05 & 60,17 & 69,68 \\ 79,40 & 76,98 & 80,29 \\ 80,51 & 66,78 & 79,33 \end{bmatrix}$$

So matriks **Y** is processed so that matriks G^2 dan H^2 is

$$G^2 = \begin{bmatrix} -0,07 & -0,02 \\ -0,44 & -0,26 \\ -0,001 & 0,10 \\ -0,16 & -0,03 \\ -0,06 & 0,08 \\ -0,06 & 0,24 \\ -0,04 & 0,06 \\ -0,25 & -0,20 \\ 0,08 & -0,08 \\ -0,17 & 0,01 \\ 0,12 & 0,05 \\ 0,25 & 0,01 \\ 0,03 & -0,21 \\ -0,06 & 0,11 \\ 0,13 & 0,26 \\ 0,65 & -0,47 \\ 0,31 & 0,24 \\ -0,22 & -0,39 \\ -0,05 & 0,48 \end{bmatrix} \qquad H^2 = \begin{bmatrix} -17,84 & 4,94 \\ -22,92 & -7,20 \\ -23,78 & 3,24 \end{bmatrix}$$
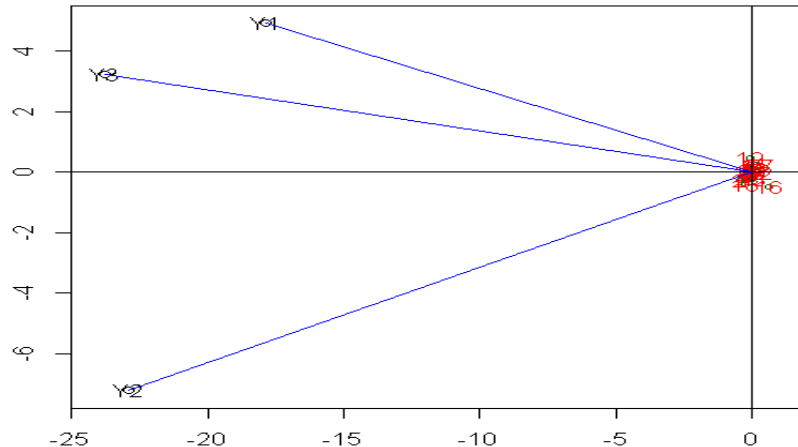
Based on G^2 dan H^2 obtain figure biplot is

Figure 2: Biplot Regency / City based on the results of the National Examination in 2016 for elementary education level

Based on the above biplot can be taken information about the diversity of variables. The Mathematics subject variable vector (Y2) has a variable vector that is longer than the other vector variables. The longer the vector variable, the greater the level of variability of the variable. This shows the results of the National Examination of students in elementary school Mathematics level in the Regency / City in West Sumatra Province is diverse. This means that there are some areas that have a relatively higher average National Examination and some areas have a relatively lower National Examination.

The IPA variable vector (Y3) forms a small angle with the Indonesian variable vector (Y1). That is, students who have a relatively high IPA National Examination Scores, tend to also have a relatively

high value on the subjects of Bahasa . However, overall all subjects tested in the National Examination of elementary level have a positive relationship with each other.
If the biplot image is enlarged then seen grouping the city / district based on similarity

Table 3: District / City Group based on the similarity of National Examination   Results of 2016 Elementary Schools

Grup	District/City
1	Bukittinggi City
2	Agam District
	Solok Selatan District
	Padang City
	Payakumbuh City
	Padang Panjang City
	Sawahlunto City
	Pariaman City
3	Solok City
	Lima Puluh Kota District
	Tanah Datar District
4	Dharmasraya District
	Sijunjung District
5	Solok District
	Pesisir Selatan District
6	Pasaman District
7	Pasaman Barat District
	Padang Pariaman District
8	Kepulauan Mentawai District

4. Conclusion

There are 9 regency / municipal groups formed based on similarities of National Examination results of junior high school in West Sumatera Province. For Basic School Level consists of 8 groups having the same characteristics. At the SMP level the English variable vector is longer than the other vector variables. This shows that English variables have the greatest diversity. While at the primary school level the greatest diversity is Mathematics. At the junior level the most powerful or positive relationship is the relationship of Mathematics and English subjects. For elementary school level IPA subjects have a strong relationship with Bahasa.

Reference

[1] BSNP dan Kalitbang Kemdikbud. 2016. *Buku Saku Ujian Nasional*. Jakarta: Departemen Pendidikan dan Kebudayaan.

[2] Gabriel, R. 1971. The Biplot Graphic Display of Matrices with Application to Principal Component Analysis. Journal of Biometrika 58: 453-467.

[3] Howard, A., dan Rorres, C. 2004. Aljabar Linier Elementer Versi Aplikasi. Jakarta: Erlangga.

[4] Mattjik, A. A., & Sumertajaya, I. M. (2011). Sidik Peubah Ganda dengan Menggunakan Aplikasi SAS. Bogor: IPB Press.

[5] Sartono, dkk. 2003. Analisis Peubah Ganda. Jurusan Statistika FMIPA IPB. Bogor

[6] Suprapto, J. 2004. *Analisis Multivariat Arti dan Interpretasi. Jakarta*: Rineka Cipta