International Journal Of Health Medicine and Current Research
E - ISSN : 2528 - 3189
P - ISSN : 2528 - 4398

International Journal of Health Medicine and Current Research
Vol. 3, Issue 01, pp.799-806, March, 2018

DOI:
10.22301/IJHMCR.2528-3189.799

Article can be accessed online on:
http://www.ijhmcr.com

ORIGINAL ARTICLE

INTERNATIONAL JOURNAL OF HEALTH MEDICINE AND CURRENT RESEARCH

CLASSIFICATION OF DATA HEALTH POLICY FOR DETERMINING HEALTH DEGREE IN MALUKU PROVINCE WITH K-MEANS METHOD

M. W. Talakua[1*], Julia Nimasratu[2], E. R. Persulessy[3], Y. A. Lesnussa[4], M. Y. Matdoan[5]

[1,2,3] Mathematics Department Faculty of Mathematics and Natural Sciences, Pattimura University
Jalan Ir. M. Putuhena, Unpatti, Poka, Ambon, Indonesia.

ARTICLE INFO

Article History:
Received 15th January, 2017
Received in revised form
16th February, 2018
Accepted 01th March, 2018
Published online 20th March, 2018

Key words:
Degree of Health, Cluster Analysis, K-Means.

Correspondence to Author:
M. W. Talakua
Mathematics Department Faculty of Mathematics and Natural Sciences, Pattimura University
Jalan Ir. M. Putuhena, Unpatti, Poka, Ambon, Indonesia

E-mail:
ocat_14@yahoo.com

ABSTRACT

One indicator to assess the success rate of a country's development can be seen from the country's achievement in providing health insurance. This study discusses the utilization of K-Means Algorithm for clustering or grouping of Regency and City in Maluku Province based on the similarity of regional characteristics in terms of indicators of mortality of local health status, namely birth rate, crude mortality rate, infant mortality rate and under-five mortality rate and mother mortality rate. The results obtained from this research, there are three groups and also there are some differences that are found according to the characteristics of each variable.

Citation: M. W. Talakua[1*], Julia Nimasratu[2], E. R. Persulessy[3], Y. A. Lesnussa[4], M. Y. Matdoan[5], 2018 "Classification Of Data Health Policy For Determining Health Degree In Maluku Province With K-Means Method", *International Journal of Health Medicine and Current Research*, 3, (01), 799-806.

INTRODUCTION

One indicator to assess the success rate of a country's development can be seen from the level of the country's achievement in providing health insurance. The government through the Ministry of Health established a number of indicators as a benchmark for the progress of health development at various levels of regional units in Indonesia, starting from the level of Province, District, to sub-district.

Data processing methods of health indicators that have been applied are still based on basic statistical techniques, such as calculations based on the average results of all indicators or based on the distribution of data. The K-

Means algorithm works by dividing data in a number of clusters to analyze similarity or dissimilarity factors attached to the data set. And then, it was analyzed pattern of relationship between data.

This study discusses the utilization of K-Means Algorithm for clustering or grouping of Regencies and Cities in Maluku Province. These clustering based on the similarity of regional characteristics in terms of mortality indicators of local health status, namely birth rate, crude mortality rate, infant mortality rate and under-five mortality rate and mortality rate mother.

METHODS

The data in this research derived from the Badan Pusat Statistik (BPS) of Maluku Province. The analysis technique used in this research is Cluster analysis with Non-Hierarchy method. There are some variable used in this research like Crude Birth Rate Variables (x_1), Crude Mortality Rate Variable (x_2), Infant Mortality Variable (x_3) and Underfive Mortality Variable (x_4) and Maternal Mortality Variable (x_5).

RESULTS AND DISCUSSIONS

Data of Health Degree Indicator

Health Degrees indicator are used like the crude birth rate (AKLK), rough mortality rate (AKMK), infant mortality (AKB), underfive mortality (AKABA), and maternal mortality (AKI). The following Table 1 is the data about health degrees indicators in several regencies and cities in Maluku Provinces.

Table 1. Data of Health Degrees Indicators.

No.	Regencies/Cities	AKLK	AKMK	AKB	AKABA	AKI
1.	Southeast West Maluku	1,990	56	82	16	8
2.	Southeast Maluku	2,423	25	32	8	9
3.	Central Maluku	7,326	86	3	1	5
4.	Buru	2,141	49	42	15	10
5.	South Buru	1,238	57	26	5	4
6.	Seram West Region	3,803	30	16	0	6
7.	Seram Eastern	2,071	31	37	18	11
8.	Aru Island	1,445	18	15	22	8
9.	Southwestern Maluku	1,405	15	4	2	0
10.	Ambon	1,550	45	18	6	2
11.	Tual	1,066	32	21	4	2

The Grouping Process of Non Hierarchy Method

The data were processed and analyzed by 11 districts / cities indicated as crude birth rate, crude mortality rate, infant mortality rate, infant mortality rate and maternal mortality rate. Table 2 below shows descriptive statistics of five variable in 11 districts / cities in Maluku province.

Table 2. *Descriptive Statisics.*

	Minimum	Maximum	Mean	Std. Deviation
Birth Rate	1	7	2	1,78
Crude Mortality Rate	15	86	40,36	20,84
Infant Mortality Rate	3	82	26,90	22,05
Underfive Mortality Rate	0	22	8,81	7,61
Maternal Mortality Rate	0	11	5,90	3,61

Quick Cluster

Table 3. *Initial Cluster Centers.*

	Cluster				
	1	2	3	4	5
Birth Rate	-0,559	-0,559	2,795	-0,559	-0,559
Crude Mortality Rate	0,750	-1,216	2,189	-1,072	0,798
Infant Mortality Rate	2,498	-1,038	-1,084	-0,540	-0,041
Underfive Mortality Rate	0,943	-0,895	-1,026	1,731	-0,501
Maternal Mortality Rate	0,577	-1,633	-2,512	0,577	-0,527

Based on Table 3 above, is the first view (initial) process of data clustering before iterating. Because later will be generated clustering process after iteration which is actually the end result of the cluster, then this output will not be analyzed.

Iteration Process

Table 4. Iteration History.

Iteration	*Change in Cluster Centers*				
	1	2	3	4	5
1	1,005	0,639	0	1,160	0,795
2	0	0	0	0	0

Based on Table 4 above, it can be seen that the iteration process is done 5 times. This process is done to get the right cluster in grouping 11 districts / cities. It is known that the minimum distance between the cluster centers that occurs from the iteration result is 2,536.

Analisis Cluster

The result of the clustering process can be seen in the table 5. below:

Table 5. Final Cluster Centers.

	Cluster				
	1	2	3	4	5
Birth Rate	-0.279	-0.559	2.759	-0.186	-0.186
Crude Mortality Rate	0.582	-0.809	2.184	-0.753	0.174
Infant Mortality Rate	1.591	-0.653	-1.084	0.049	-0.313
Underfive Mortality Rate	0.877	-0.764	-1.026	0.943	-0.676
Maternal Mortality Rate	0.854	-1.356	-0.251	0.946	-0.527

In Table 5 above, it can be seen that the data above are still related to the standardization process which refers to z-score with the following conditions:

1. A negative number (-) means the data is below a total average
2. A positive number (+) means the data is above the total average

Cluster Formulas

If applied to the variable of Crude Birth Rate:

a. Average Crude Birth Rate in cluster 1: (Average Crude Birth Rate) - (0.279 x Standard Deviation of Crude Birth Rate).

b. Average Crude Birth Rate in cluster 2: (Average Crude Birth Rate) - (0.559 x Standard Deviation of Crude Birth Rate).

c. Average Crude Birth Rate in cluster 3: (Average Crude Birthrate) + (2,795 x Standard Deviation of Crude Birth Rate).

d. Average Crude Birth Rate in cluster 4: (Average Crude Birth Rate) - (0.186 x Standard Deviation of Crude Birth Rate).

e. Average Crude Birth Rate in cluster 5: (Average Crude Birthrate) + (0.186 x Standard Deviation of Crude Birth Rate).

And so on for the interpretation of other data. The average of crude birth rate and standard deviation can be seen in Table 4.2 while the average population can be seen in Table 4.5 above.

Interpretation of Cluster Numbers

From the cluster process, there are 5 clusters or 5 groups with each characteristic that are different from each other. Differences can be traced per variable, based on the interpretation of the sign (+) and the sign (-) and the number itself.

Before interpreting the contents of each cluster, first we show an example of interpretation of a variable ZAKLK. Because the number from cluster 1 and cluster 2 are negative while in the cluster 3 is positive numbers, then it can be said the average Crude Birth Rate on cluster 3 is higher than the average Crude Birth Rate on cluster 1 and cluster 2, and the average Crude Birth Rate of cluster 4 and cluster 5 is lower than cluster 3.

Cluster Formula

By using this formula can calculate the value of each variable in each cluster.

$$X = \mu + Z \cdot \sigma$$

with :

X = The average sample (average variable on each cluster)

μ = The average population

σ = Standart Deviation

Z = The value of standarisation in SPSS

By using a formula to calculate the average Crude Birth Rate on each cluster, can be derived:

1. The total average Crude Birth Rate is 2,000
2. The total standard deviation of Crude Birth Rate is 1.78885

Therefore:

a. Average Crude Birth Rate of cluster 1 = 2,0000 + (-0,27951 x 1,78885) = 1,499998537 atau 4,999
a. Average Crude Birth Rate of cluster 2 = 2,0000 + (-0,55902 x 1,78885) = 0,999997073atau 0,999
b. Average Crude Birth Rate of cluster 3 = 2,0000 + (2,79508 x 1,78885) = 6,999978858 atau 6,999
c. Average Crude Birth Rate of cluster 4 = 2,0000 + (-0,18634 x 1,78885) = 1,666665691 atau 1,666
d. Average Crude Birth Rate of cluster 5 = 2,0000 + (-0,18634 x 1,78885) = 1,666665691 atau 1,666

To calculate the average Crude Mortality Rate on each cluster, can be done with the formula above, and can be derived:

1. The average of total Crude Mortality Rate is 40.3636
2. The standard deviation of total Crude Mortality Rate is 20.84357

Therefore:

a. Average Crude Mortality Rate in Cluster 1 = 40,3636 + (0,58226 x 20,84357) = 52,49997707 atau 52,499
a. Average Crude Mortality Rate in Cluster 2 = 40,3636 + (-0,80906 x 20,84357) = 23,49990126 atau 23,499
b. Average Crude Mortality Rate in Cluster 3 = 40,3636 + (2,18947 x 20,84357) = 85,99997121 atau 85,999
c. Average Crude Mortality Rate in Cluster 4 = 40,3636 + (-0,75308 x 20,84357) = 24,6667243 atau 24,666
d. Average Crude Mortality Rate in Cluster 5 = 40,3636 + (-0,17446 x 20,84357) = 36,72723078 atau 36,727

To calculate the average Infant Mortality Rate on each cluster, can be done with the formula above.

1. The total average infant mortality rate is 26.9091
2. The total standard deviation of infant mortality rate is 22.05200

Therefore:

a. Average Infant Mortality Rate in Cluster 1 = 26,9091 + (1,59128 x 22,05200) = 62,00000656 atau 62,000
b. Average Infant Mortality Rate in Cluster 2 = 26,9091 + (-0,65341 x 22,05200) = 12,500010268 atau 12,500
c. Average Infant Mortality Rate in Cluster 3 = 26,9091 + (-1,08421 x 22,05200) = 3,00010108 atau 3,000
d. Average Infant Mortality Rate in Cluster 4 = 26,9091 + (0,04947 x 22,05200) = 28,00001244 atau 28,000
e. Average Infant Mortality Rate in Cluster 5 = 26,9091 + (-0,67664 x 22,05200) = 11,98783472 atau 11,987

To calculate the average of Underfive Mortality Rate on each cluster, can be done with the formula above.

1. The total average of Underfive Mortality Rate is 8.8182
2. The total standard deviation of Underfives Mortality Rate is 7.61339

Therefore:

a. Average Underfive Mortality Rate in Cluster 1 = 8,8182 + (0,87764 x 7,61339) = 15,5000156 atau 15,000
b. Average Underfive Mortality Rate in Cluster 2 = 8,8182 + (-0,76420 x 7,61339) = 3,000047362 atau 3,000
c. Average Underfive Mortality Rate in Cluster 3 = 8,8182 + (-1,02690 x 7,61339) = 1,000009809 atau 1,000
d. Average Underfive Mortality Rate in Cluster 4 = 8,8182 + (0,94331 x 7,61339) = 15,99998692 atau 15,999
e. Average Underfive Mortality Rate in Cluster 5 = 8,8182 + (-0,52764 x 7,61339) = 4,8010709 atau 4,801

To calculate the average maternal mortality rate in each cluster, can be done with the formula above.

1. The total of average maternal mortality rate is 5.9091
2. The total of standard deviation of maternal mortality is 3.61814

Therefore:

a. Average Maternal Mortality Rate in cluster 1 = 5,9091 + (0,85428 x 3,61814) = 9,00004639 atau 9,000
b. Average Maternal Mortality Rate in cluster 2 = 5,9091 + (-1,35680 x 3,61814) = 1,000007648atau 1,000
c. Average Maternal Mortality Rate in cluster 3 = 5,9091 + (-0,25126 x 3,61814) = 5,000006144 atau 5,000
d. Average Maternal Mortality Rate in cluster 4 = 5,9091 + (0,94641 x 3,61814) = 9,333705691 atau 9,333
e. Average Maternal Mortality Rate in cluster 5 = 5,9091 + (-0,52764 x 3,61814) = 4,00002461 atau 4,000

Interpretation of Each Cluster

Based on Table 5 above, can be defined as the following:

Cluster 1 :

In cluster 1 it is stated that in 2014 the crude birthe rate above the average in 11 regencies /cities in in Maluku Province while the Crude mortality rate, the Infant Mortality Rate, the Underfive Mortality Rate and the Maternal Mortality Rate are below average. Based on the above characteristics, it can be said that the Crude Birth Rate is very minimal while the number of Crude mortality rate, Infant mortality rate, Underfive mortality rate and Maternal mortality rate is inversely proportional to the Crude birth rate which is very much happening in 2014.

Cluster 2 :

In cluster 2 it is stated that in 2014 the number of Crude birth rate, Crude mortality rate, Infant mortality rate, Underfive mortality rate and Maternal Mortality rate are below the average in 11 regencies / cities in Maluku Province. From the above characteristics, it can be said that the Crude birth rate, Crude mortality rate, Infant mortality rate, Underfive mortality rate and Maternal Mortality Rate are less occur in 2014.

Cluster 3 :

In cluster 3 it is stated that in the year 2014 the number of Crude birth rate and Crude mortality rate are above the average, while Infant mortality rate, Underfive mortality rate and Maternal mortality rate are below average in 11 regencies/cities in Maluku Province. From the characteristics above, it can be said that the Crude birth rate and Crude mortality rate are less occur while the happening of Infant mortality rate, Underfive mortality rate and Maternal mortality rate are more occur in 2014.

Cluster 4 :

In cluster 4 it is declare that in the year 2014 the Crude birth rate and Crude mortality rate are above the average, while the Infant mortality rate, Underfive mortality rate and Maternal mortality rate are below the average in 11 regencies/cities in Maluku Province. From the characteristics above, it can be said that the number of Crude birth rate and Crude mortality rate are more occur, while Infant mortality rate, Underfive mortality rate and Maternal mortality rate are less occur in 2014. This is inversely proportional to cluster 3 above.

Cluster 5 :

In cluster 5 it is states that in 2014 the Crude mortality rate is below the average, while the number of Crude birth rate, Infant mortality rate, Underfive mortality rate and Maternal mortality rate are above the average in 11 regencies/cities in Maluku Province. From the characteristics above, it can be said that Crude mortality rate is more occur, while the number of Crude birth rate, Infant mortality rate, Underfive mortality rate and Maternal mortality rate are less occur in 2014.

Differences Variable on Formed Clusters

After forming 5 clusters, the next step is to see whether the variables that have formed the cluster have differences on each cluster, in this case can be seen from the value of F and the probability value (Sig) of each variable. This is done by looking at the following ANOVA output, like the following:

Table 6. Anova.

	Cluster		Error		F	Sig
	Mean Square	df	*Mean Square*	df		
Birth Rate	2,201	4	0,200	6	11,022	0,006
Crude Mortality Rate	2,143	4	0,238	6	9,017	0,010
Infant Mortality Rate	1,849	4	0,434	6	4,259	0,057
Underfive Mortality Rate	1,952	4	0,366	6	5,337	0,035
Maternal Mortality Rate	2,184	4	0,212	6	10,282	0,007

Based on Table 6 this can be seen in the cluster column shows the size between cluster mean, while the error column shows the size within cluster mean, so column F is:

$$F = \frac{Between\ Means}{Within\ Means}$$

1. The F value on ZAKLK is derived from:

$$F = \frac{Between\ Means\ ZAKHK}{Within\ Means\ ZAKHK} = \frac{2,201}{0,200} = 11,022$$

2. The F value on ZAKMK is derived from:

$$F = \frac{Between\ Means\ ZAKMK}{Within\ Means\ ZAKMK} = \frac{2,143}{0,238} = 9,017$$

3. The F value on ZAKB is derived from:

$$F = \frac{Between\ Means\ ZAKB}{Within\ Means\ ZAKB} = \frac{1,849}{0,434} = 4,259$$

4. The F value on ZAKABA is derived from:

$$F = \frac{Between\ Means\ ZAKABA}{Within\ Means\ ZAKABA} = \frac{1,952}{0,366} = 5,337$$

5. The F value on ZAKI is derived from:

$$F = \frac{Between\ Means\ ZAKI}{Within\ Means\ ZAKI} = \frac{2,182}{0,212} = 10,282$$

Interpretation F Value and the Significant:

In principle, the greater number of F variable, and the significant number is 0.05, then the greater difference of the variables in five variables. For example, the largest F number (11,022) in ZAKLK, with the SIG column number is 0.006, meaning the significance is real. This means that the Crude birth rate factor greatly differences the characteristics of the five clusters, or it could be said that the Crude birth rate on the five clusters is very different between clusters one and other clusters. Note the difference with the ZAKABA variable, which has the value F (5.337) and SIG 0.035. Because the significant numbers are still below 0.05 (0.035 <0.05), then the ZAKABA variable in cluster 1, cluster 2, cluster 3, cluster 4 and cluster 5 still have differences. However, the value of F ZAKLK variable is greater than the value of F ZAKABA variable, and it can be interpreted the Crude birth rate more than the Underfive birth rate on the five clusters.

Tabel 7. *Number of cases in each cluster.*

Cluster	1	2
	2	2
	3	1
	4	3
	5	3
Valid		11
Missing		0

Based on Table 4.7 it can be seen that the most data are in cluster 4 and cluster 5 i.e. 3 regencies / cities while the least data is in cluster 3 i.e. 1 regency/city, without missing variables. Therefore, all data of 11 regencies/ cities are complete maping on the three cluser, with the composition above, because cluster 4 and cluster 5 is the largest cluster.

Analisys Cluster Composition

Table 8. *Cluster Membership.*

No.	Regency/City	Cluster	Distance
1.	Southeast West Maluku	1	1,005
2.	Southeast Maluku	4	1,087
3.	Central Maluku	3	0,000
4.	Buru	1	1,005
5.	South Buru	5	0,795
6.	Seram West Region	5	1,256
7.	Seram Eastern	4	0,758
8.	Aru Island	4	1,160
9.	Southwestern Maluku	2	0,639
10.	Ambon	5	0,741
11.	Tual	2	0,639

a. Cluster 1: it can be seen that Southeast West Maluku and Buru is less occurrence of crude birth rate but often occurrence of Crude Death Rate, Infant Mortality Rate, Underfive Mortality Rate and Maternal Mortality Rate in 2014.

b. Cluster 2: Southwestern Maluku and Tual is often occur the Crude birth rate, Crude Mortality rate, Underfive mortality rate and Maternal mortality rate in 2014

c. Cluster 3: In Central Maluku Regency there are many cases of Crude birth rate and Crude mortality rate but less occur Infant mortality rate, Underfive mortality rate and Maternal mortality rate in 2014.

d. Cluster 4: In Southeast Maluku, Seram Eastern and Kep. Aru are less occurence of Crude birth rate and Crude mortality rate. However, it is often occur the Infant mortality rate, Underfive mortality rate and Maternal mortality rate in 2014.

e. Cluster 5: South Buru, Seram Barat and Ambon City are often occurence of Crude mortality rate, but less occur of Crude birth rate, Infant mortality rate, Underfive mortality rate and Maternal mortality rate in 2014.

Conclusion

The grouping of regencies/cities based on Crude birth rate, Crude mortality rate, infant mortality rate, Underfive mortality rate and Maternal mortality rate in 2014 divided in 5 clusters, like the following:

a. Cluster 1: it can be seen that Southeast West Maluku and Buru is less occuring of Crude birth rate, but often occur of Crude mortality rate, Infant mortality rate, Underfive mortality rate and Maternal mortality rate in 2014.

b. Cluster 2: In Southwestern Maluku and Tual City are often occuring the Crude birth rate, Crude mortality rate, Infant mortality rate, Underfive mortality rate and Maternal mortality rate in 2014.

c. Cluster 3: In Central Maluku there are many cases of Crude birth rate and Crude mortality rate, but fewer cases of Infant mortality rate, Underfive mortality rate and maternal mortality rate in 2014.

d. Cluster 4: In Southeast Maluku, Seram Eastern and Kep. Aru are less occur of Crude birth rate and Crude mortality rate, but often occurring the number of infant mortality rate, Underfive mortality rate and Maternal mortality rate in 2014.
e. Cluster 5: In South Buru, Seram West Region and Ambon City are often occur of Crude mortality rate but lack occur of Crude birth rate, Infant mortality rate, Underfive mortality rate and Maternal mortality rate in 2014.

REFERENCES

1. Asmalaizza. Pemerataan Akses Pendidikan Bagi Masyarakat Miskin dan Terpencil 2009.
2. Cahyawati, D. Karakteristik Anak Putus Sekolah Pendidikan Dasar Menggunakan Metode Chi Square Automatic Interaction Detection (CHAID. Jurnal Penelitian Sains. Fakultas MIPA. Universitas Sriwijaya. Palembang; 2007.
3. Cahyawati, D., Pemodelan Masalah Anak Putus Sekolah Pendidikan Dasar. Jurnal Ilmiah MIPA. Fakultas MIPA. Universitas Lampung. Lampung; 2007.
4. Cahyawati, D., Dwipurwani O., Agustien M., Pendekatan Model Risiko Putus Sekolah Pendidikan Dasar pada Masyarakat Miskin sebagai Upaya Pemerataan Akses Pendidikan di Kabupaten Ogan Ilir menggunakan fungsi pengelompokkan analisis diskriminan. Laporan Penelitian. Fakultas MIPA. Universitas Sriwijaya. Palembang; 2010.
5. Cahyawati D, Dwipurwani O., Pemetaan Biplot untuk Masalah Putus Sekolah Pendidikan Dasar pada Masyarakat Miskin antar Kecamatan di Kabupaten Ogan Ilir. Fakultas MIPA. Universitas Sriwijaya. Palembang. Jurnal Penelitian Sains. 2011;14(2).
6. Tatham, R. L. , Hair, J. F, Anderson, R. E. , dan Black, W. C., "Multivariate Data Analysis", prentice Hall, New Jersey; 1998.
7. Vinsent Gaspersz, Fungsi Diskriminan (Diskriminant Function); 1995.

****