CLUSTER ANALYSIS OF COVID-19 IMPACT ON POVERTY IN INDONESIA USING SELF-ORGANIZING MAP ALGORITHM

Ika Nur Laily Fitriana[1], Mohammad Okky Mabruri[2]

[1]Departement of Statistics, Institut Teknologi Sepuluh Nopember
[2]Idata1011 (Data Science Community)
e-mail: [1]ika.206003@mhs.its.ac.id, [2]okkymabrur@gmail.com

Abstract

The increase in poverty rates caused by the COVID-19 pandemic requires immediate attention from policymakers. Each province in Indonesia has unique characteristics of poverty, and as a result, each province's response to COVID-19's impact on poverty is unique. As a result, a provincial cluster analysis based on the similarity of poverty characteristics is necessary to identify provinces that require increased vigilance. The purpose of this study is to cluster Indonesian provinces according to their similarity in terms of poverty impact before and during COVID-19. The impact of poverty prior to and during COVID-19 is quantified by comparing 2021 (during COVID-19) to 2019 (before COVID-19). We discovered that the COVID-19 has a significant impact on poverty. Hybrid SOM-Kmeans with three clusters is the optimal method for producing the smallest Davies-Bouldin Index. COVID-19 has high, moderate, and low impact on poverty, respectively. Cluster 1 is a cluster with a significant impact on poverty in a province where tourism is the primary industry. Due to sluggish tourism, the community's purchasing power is diminished, thereby increasing poverty. Cluster 3, namely Papua, has a low impact due to its primary sector characteristics in the mining sector.

Abstrak

Peningkatan angka kemiskinan akibat pandemi COVID-19 membutuhkan perhatian segera dari pengambil kebijakan. Setiap provinsi di Indonesia memiliki karakteristik kemiskinan yang berbeda, sehingga penanggulangan setiap provinsi terhadap dampak COVID-19 terhadap kemiskinan pun berbeda. Oleh karena itu, analisis klaster provinsi berdasarkan kesamaan karakteristik kemiskinan diperlukan untuk mengidentifikasi provinsi yang memerlukan perhatian lebih. Tujuan dari penelitian ini adalah untuk mengelompokkan provinsi-provinsi di Indonesia menurut kesamaannya dalam hal dampak kemiskinan sebelum dan selama COVID-19. Dampak kemiskinan sebelum dan selama COVID-19 diukur dengan membandingkan data tahun 2021 (selama COVID-19) dengan data tahun 2019 (sebelum COVID-19). Penelitian ini menemukan bahwa COVID-19 memiliki dampak signifikan terhadap kemiskinan. Metode Hybrid SOM-Kmeans dengan tiga cluster merupakan metode yang paling optimal karena menghasilkan Davies-Bouldin Index terkecil. Pembagian tiga klaster berdasarkan dampak COVID-19 terhadap kemiskinan yaitu tinggi, sedang, dan rendah. Klaster 1 merupakan provinsi yang terkena dampak paling tinggi karena sektor utama pendapatan provinsi adalah pariwisata. Akibat lesunya pariwisata, daya beli masyarakat berkurang, sehingga meningkatkan kemiskinan. Klaster 3 yaitu Papua yang terdampak rendah karena karakteristik sektor utamanya di sektor pertambangan.

INTRODUCTION

The COVID-19 pandemic and measures to prevent its spread have resulted in a severe global economic contraction. According to the International Monetary Fund, the COVID-19 pandemic has resulted in a global recession, increasing unemployment and poverty in every country. Due to the adoption of different regulations aimed at limiting the spread of COVID-19 or attempts to break the chain of its dissemination, many economic operations have shrunk and even stopped production. As a consequence, unemployment has increased, individual and corporate productivity has decreased, and Indonesia's poverty rate has increased. [1]. The number of poor people has been increasing since the COVID-19 pandemic, even though before COVID-19, it showed a downward trend. Between March 2015 and September 2019, poor people consistently decreased from 28.59 million to 24.79 million. In March 2018, Indonesia had approximately 25.95 million poor people, which decreased by 280 thousand people in September 2018, 530 thousand people until March 2019, and 350 thousand people in September 2019. Meanwhile, between September 2019 and March 2021, the number and prevalence of poverty increased in urban, rural, and national areas. In March 2021, the number of poor people reached 27.54 million, increasing 1.12 million (4.23 percent) over March 2019. Relevant to the pandemic-induced increase in the number and prevalence of poverty, the average expenditure of the poor tends to be further away from the poverty line, as evidenced by the poor vulnerable groups' (labor and informal sector workers) descent into poverty the poor becoming poorer. The COVID-19 pandemic exacerbates and deepens poverty [2].

The enhancement in poverty rates resulting from the COVID-19 pandemic needs serious attention from policymakers at the national and regional levels. The government must make concerted efforts to alleviate poverty in the short, medium, and long term [3]. Each province in Indonesia has different characteristics of poverty, and thus each province's response to the impact of COVID-19 on poverty is different. As a result, a provincial cluster analysis based on the similarity of poverty characteristics is required to understand the provincial groups that require increased vigilance. Cluster analysis is a technique for grouping data objects based on similar variables or characteristics [4]. Data objects with a high similarity will be clustered together, while those with a low similarity or a large difference will be clustered separately [5]. Several studies related to poverty cluster analysis in Indonesia were conducted by Ferezagia [6] using the non-hierarchical cluster method to group the poverty level into three groups, namely high, medium, and low poverty. Sano and Nindito [7] researched the application of the k-means algorithm for cluster analysis of provincial poverty in Indonesia. Bahauddin, Fatmawati and Sari [8] uses the k-means algorithm to group provinces in Indonesia based on poverty levels, and the result is that there are 3 clusters, namely low poverty level, moderate poverty level, and high poverty level.

This study aims to group provinces in Indonesia based on the similarity of poverty impact before and during COVID-19. We chose the province as the research unit because the availability of data at the provincial level is more complete, in addition, the province is more representative of the condition of the entire territory of Indonesia. The impact of poverty before and during COVID-19 is measured by examining the difference between 2021 (during COVID-19) and 2019 (before COVID-19). The clustering results are expected to shed light on which provinces require additional attention regarding COVID-19's impact on poverty. The clustering method used is the Self-Organizing Map (SOM). Self-Organizing Map (SOM) is a efficient algorithm for visualizing high-dimensional data by reducing its dimensions from an n-dimensional input to a lower dimension while maintaining the original topological relationship [9]. Additionally, the SOM algorithm is unique because it combines the objectives of projection and clustering algorithms. It can be used to visualize the clusters in a data set while also representing it on a two-dimensional map to preserve the nonlinear relationships between the data items; nearby items are located close together on the map [10].

MATERIALS AND METHODS

This section will briefly describe the material and methods conducted in this study. The material is data, and the methods are different test and SOM.

1. Dataset

This study used secondary data from the Central Bureau of Statistics (BPS) Indonesia. The data is the first semester of 2019 (March) and the first semester of 2021 (March). Data for 2019 show conditions before COVID-19 and data for 2021 shows conditions during COVID-19. The sample unit is a province in Indonesia, which consists of 34 provinces. Research variables in this study are shown in Table 1.

2. Methods

The impact of COVID-19 pandemic on poverty is explored in this study using cluster analysis. As an application of the variables used, the 2019 data will be compared with the data in 2021. First, we use a nonparametric location test to test whether there is a significant difference before COVID-19 (data in 2019) and during COVID-19 (data in 2021). Then we calculate the difference between 2019 and 2021 data by subtracting the 2021 data from 2019. The data difference will be used for cluster analysis. Using the difference data is to group provinces in Indonesia based on changes in data from 2019 and 2021, whether there is a contraction or resilience in poverty due to COVID-19. We also performed a comparison of the standard SOM and Hybrid SOM methods. The Davies-Bouldin Index evaluates the best method, then interpretation and mapping of cluster results are carried out. The explanation of each method used is as follows.

Wilcoxon Signed-Ranked Test.

In the dearth of a multivariate normal distribution, a nonparametric location test can be used to test for differences in location parameters. The Wilcoxon signed-rank test is used in this study. The Wilcoxon signed-rank test is a nonparametric statistic that is used to determine the significance of a difference between two groups of paired data on an ordinal or interval scale when the data are not normally distributed [11]. The Wilcoxon signed-rank test is a substitute for the paired t-test or t paired if the data do not

Table 1. Research variables

Symbols	Variables	Description of Variables
X_1	Headcount Index	The percentage of impoverished individuals who live in poverty. The Headcount Index calculates the percentage of people who are classified as impoverished.
X_2	Poverty line	The minimal amount of rupiah required to satisfy the basic requirements of food (2100 kcal per capita per day) and non-food basic needs.
X_3	Food Poverty Line	The cost of meeting basic dietary requirements, which amounts to 2100 kcal per capita per day. There are 52 different kinds of commodities that make up commodity packages for basic food requirements.
X_4	Non-Food Poverty Line	The bare minimum in terms of shelter, clothes, education, and health. There are 51 kinds of commodities in urban areas and 47 types of commodities in rural areas in commodity packages for non-food basic requirements.
X_5	Poverty Gap Index	The average measure of each impoverished population's spending gap against the poverty line.
X_6	Poverty Severity Index	An index that provides information on the distribution of expenditure among the poor.

meet the normality assumption. The null hypothesis is that the median of differences is zero, and the alternative hypothesis is that the median of differences is not zero. If the probability (p-value) is less than or equal to 0.05, the null hypothesis is rejected. If the probability (p-value) is greater than or equal to 0.05, the null hypothesis is accepted.

Self-Organizing Map (SOM)

Self-Organizing Map (SOM) is one of the artificial neural network models that use unsupervised learning methods with the aim of grouping data. The most important feature of SOM is that it summarizes data by comparing clusters and each cluster is projected onto a map node [12]. SOM basic algorithm can be summarized as follows [13].

1. Initialization. Choose the dimension and size of the output space.
2. Sampling. Randomly select an input vector *x(t)* from the training data set.
3. Matching of similar items. Calculate the Euclidean distances between the input vector and the weight vectors of each output node. Find the best matching node c(t) at iteration t by applying the minimum distance criterion like equation (1).

$$c(t) = \arg\min_{t}\left\{\left\| x(t) - w_i(t) \right\|\right\}; i = 1, 2, \ldots, n \quad (1)$$

4. Updating the weight. Using the equation (2), adjust the weights of the winner node and its neighbors based on their distances from the winning node. The neighborhood function hci(t) will be equal to 1 for the winning node.

$$w_i(t+1) = w_i(t) + \alpha(t) h_{ci}(t)\left[x(t) - w_i(t)\right] \quad (2)$$

5. Adjust the parameter. Set *t = t +1*. Change the size of the neighborhood and the learning rate.
6. Continuation. Return to Step 2 until the weight change is less than a predetermined threshold value or the maximum number T of iterations has been achieved. Otherwise, come to a halt.

Cluster Evaluation

Clustering evaluation is carried out to know how good the quality of the clustering results is. In this study, the evaluation of clustering results used was the Davies-Bouldin Index. Davies-Bouldin Index is a method that aims to evaluate clusters in a clustering method based on the maximum distance between clusters (inter clusters) and distance between points (intra values). The smaller the DB index value, the better the cluster results [14]. The formula of Davies-Bouldin Index is defined in equation (3).

$$DB = \frac{1}{K}\sum_{i=1}^{K} R_i \quad (3)$$

where K is the number of clusters. The formula of R_i is shown in equation (4).

$$R_i = \max_{j \neq i} \frac{S_i + S_j}{d_{ij}} \quad (4)$$

where S_i and S_j are point to the spread of within clusters for *i*-th and *j*-th [15].

RESULTS AND DISCUSSIONS

1. Exploratory Data Analysis

Before conducting cluster analysis, we exhibit the descriptive statistics of each variable in 2019 and 2021, that shown in Table 2. In general, it can be concluded that there is an increase in poverty from 2019 to 2021. The average percentage increase for each poverty variable can be seen in Figure 1. The percentage increase in Figure 1 is obtained from the average difference between 2021 and 2019. The highest average percentage increase was in the Poverty Severity Index (X_6), which was 14.6%. The higher the Poverty Severity Index, the higher of disparity in spending among the poor. The higher Poverty Severity Index has implications for the growing inequality in spending for the necessities of life, both primary and secondary. In connection with the COVID-19 pandemic, poverty development (amount, depth, severity) is strongly influenced by income and the Poverty Line. Reductions in income due to reduced working time, job loss, and loss of business have a more significant impact on poverty. Due to declining incomes, purchasing

Table 2. Descriptive statistics of variables

Variable	Year	Mean	Standard Deviation	Minimum	Q1	Median	Q3	Maximum
X_1	2019	10.46	5.68	3.47	6.30	8.76	13.75	27.53
	2021	10.76	5.40	4.53	6.65	8.90	13.29	26.86
X_2	2019	460,090	95,322	327,402	385,869	442,843	504,934	677,717
	2021	513,843	103,283	364,251	426,428	501,614	573,460	752,203
X_3	2019	341,098	64,859	245,761	286,707	338,761	383,076	492,693
	2021	380,825	69,694	279,240	318,311	379,733	429,885	544,017
X_4	2019	118,992	34,610	73,626	97,656	109,816	131,540	207,345
	2021	133,018	37,132	84,504	106,117	125,824	148,673	218,306
X_5	2019	1.84	1.44	0.50	0.91	1.41	2.33	7.17
	2021	1.90	1.25	0.61	1.04	1.51	2.42	5.60
X_6	2019	0.48	0.52	0.10	0.21	0.31	0.58	2.60
	2021	0.50	0.41	0.11	0.24	0.38	0.63	1.96

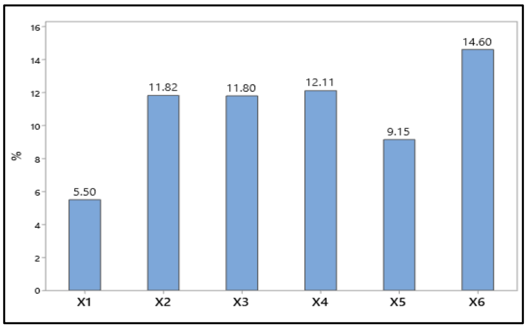


Figure 1. Bar chart of percentage of average difference in poverty variables

power weakens. Moreover, people need medical expenses and additional vitamin consumption to have a body immune that can withstand the effects of COVID-19. The lowest average percentage increase is found in the Headcount Index (X_1), which is 5.50%.

2. Nonparametric Different Test

This section provides multivariate difference test analysis for all variables in 2019 and 2021 using hypothesis testing. We perform hypothesis testing to verify a significant difference between poverty data in 2019 and 2021. We perform multivariate normality tests before doing difference tests. The multivariate normality test for all variables in 2019 and 2021 indicates that the data do not follow the multivariate normal distribution, necessitating the use of a nonparametric method, namely the Wilcoxon Signed-Ranked Test. The results of the normal multivariate test indicated that the data did not conform to the normal multivariate distribution assumption. Using the multivariate nonparametric location test, we determined that the median of differences is not equal to zero, with a statistic of 5.086 and a p-value of 0.0001. The results of univariate nonparametric location tests are shown in Table 2. Each variable is significantly different from the others. Each variable has a p-value less than the level of significance (0.05). A negative sign in estimates value indicates an increase between 2019 and 2021.

Table 3. Result of nonparametric different test.

Variables	Estimates	Statistics	P-value
X_1	-0.23	-3.052	0.002
X_2	-53278.50	-5.086	0
X_3	-39718.50	-5.086	0
X_4	-13574	-5.086	0
X_5	-0.08	-2.359	0.017
X_6	-0.025	-2.308	0.018

3. Cluster Analysis

In this sub-chapter, the analysis of Indonesia's provincial clusters based on the data difference for 2021 and 2019 based on the variables in Table 1 will be discussed. The number of clusters formed is 3 clusters representing provinces experiencing high, medium, and low impacts due to the COVID-19 pandemic in poverty. The reason for forming 3 clusters is following previous research from Ferezagia [6] and Bahauddin et.al [8], which stated that the formation of 3 clusters on poverty data was the most optimal cluster. Before being proceed, the data was scaled to 0 to 1, because the data unit is difference. We conducted cluster analysis using standard SOM algorithm comparing with Hybrid SOM-Average linkage, Hybrid SOM-Single linkage, Hybrid SOM-Complete linkage and Hybrid SOM-K means. This aims to see whether a more complex model will give the best grouping results. In standard SOM algorithm, we use 3x1 grid, because the number cluster that will be formed is 3. Meanwhile, in Hybrid method, the determination of the grid is set at 5x5 because it is sufficient to represent the data under study. Grid size that is too large will affect the unrepresentative of the data in the cluster. In addition, we also compared the performance of rectangular and hexagonal SOM topologies. Determination of the best cluster method is the Davies-Bouldin Index. The smaller the Davies-Bouldin Index, the better the cluster method. The comparison of the performance of the cluster results is shown in Table 4.

Table 4. Comparison SOM method with Davies-Bouldin Index

Methods	Hexagonal	Rectangular
SOM	0.873909	0.873909
SOM-Average Linkage	1.194451	1.291886
SOM-Single Linkage	1.194451	1.291886
SOM-Complete Linkage	1.194451	1.291886
SOM-Kmeans	1.045959	0.74847[a]

[a]*The lowest value*

The comparison of the clustering methods in Table 4 can be concluded that the best method is Hybrid SOM-Kmeans. Therefore, in the following discussion is a detailed analysis of Hybrid SOM-KMeans. The SOM network is used to divide the input pattern into several clusters. The SOM network requires training progress to minimize the average distance of an object to the nearest unit [16]. The results of the training progress are shown in Figure 2. In this study, 1000 iterations were carried out. Figure 2 explains the number of training progress which shows that the number of iterations is carried out for distances close to the closest average to get convergent results. Convergent iteration is the number of processes that the software does to get stable results. In the 700th iteration of

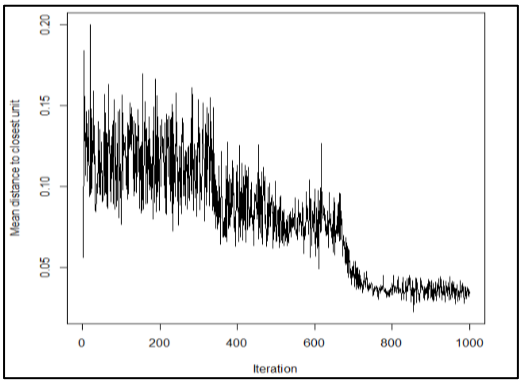
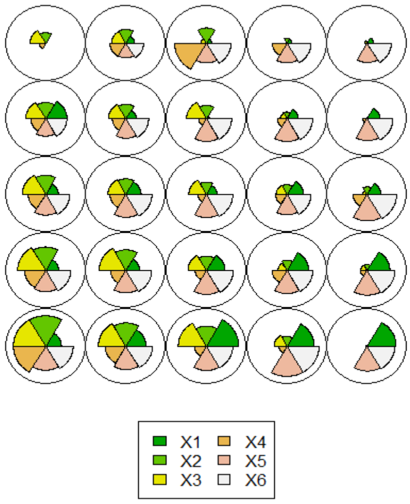
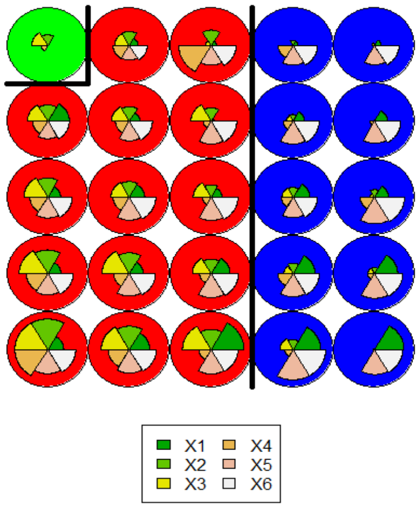


Figure 2. Training progress, as measured by the average distance of an object with the closest codebook vector unit



(a)



(b)

Figure 3. Fan diagram (a) before divided to 3 cluster, (b) after divided to 3 cluster

training progress has shown convergence. The mean distance to the closest value starts to stabilize when it is below 0.05. If the curve shows a plot that has not converged, then more iterations are needed.

The SOM algorithm process generates a SOM model and produces a diagram containing several circles called a fan diagram. Topologically the circles in the fan diagram will be close together if the characteristics are the same. The fan diagram shows the distribution of the variables. The longer the radius and the more dominant the colour of the variable indicates that the variable is superior to other variables that show the same characteristics. Figure 3 is a fan diagram using a rectangular topology with a 5x5 grid. Based on Figure 3 (b) above, the process of understanding the diagram in the SOM algorithm is when the diagram already has a colour and is bounded by vectors visualized in the mapping plot. The model formed by the SOM algorithm is then formed into 3 clusters using the k-means cluster method. Each cluster formed has its characteristics. The first cluster is marked with red nodes where the characteristics of the cluster are dominated by Poverty line (X_2), Food Poverty Line (X_3), and Non-Food Poverty Line (X_4). The second cluster is marked with blue nodes dominated by Headcount Index (X_1), Poverty Gap Index (X_5), and Poverty Severity Index (X_6). At

Table 4. Number and provincial members of each cluster using Hybrid SOM-Kmeans

Cluster	Number of Members	Cluster Members
1	16	North Maluku, West Sulawesi, Southeast Sulawesi, South Sulawesi, North Sulawesi, South Sumatra, Bengkulu, West Kalimantan, East Nusa Tenggara, Riau islands, West Nusa Tenggara, West Java, Central Java, Yogyakarta, East Java, Bali
2	17	Aceh, North Sumatra, West Sumatra, Riau, Jambi, West Papua, Maluku, Lampung, Bangka Belitung Islands, Gorontalo, DKI Jakarta, Central Sulawesi, North Kalimantan, East Kalimantan, South Kalimantan, Banten, Central Kalimantan
3	1	Papua

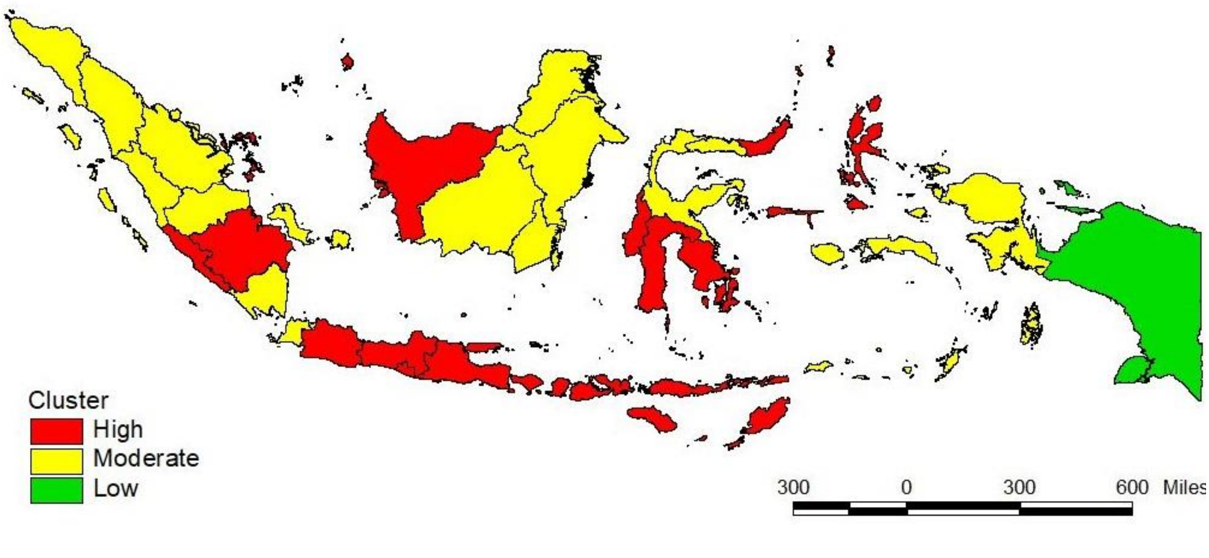


Figure 4. Provincial Mapping

Table 5. Number and provincial members of each cluster using Hybrid SOM-Kmeans

	X_1	X_2	X_3	X_4	X_5	X_6
Cluster 1 (High)	0.31	63,378.29	46,363.06	17,015.24	0.09	0.03
Cluster 2 (Moderate)	0.37	43,644.88	32,599.75	11,045.13	0.12	0.05
Cluster 3 (Low)	-0.67	51,860.00	40,976.00	10,884.00	-1.57	-0.91

the same time, the third cluster, marked with green nodes, has little influence from Headcount Index (X_1), Food Poverty Line (X_3), and Poverty Gap Index (X_5). Table 6 contains information about the number and provincial members of each cluster using Hybrid SOM-Kmeans. Figure 4 depicts the provincial map.

After knowing members of each cluster, cluster profiling is carried out by calculating the average of data in each cluster. Data used in this cluster analysis is difference data, so the extent of the impact or change in poverty can be seen before and during the COVID-19 pandemic. Profiling aims to determine the characteristics of each cluster based on changes or differences in the poverty variable in 2019 and 2021.

The characteristics of each cluster can be seen in Table 5. The characteristics of the cluster can be seen from the average difference in data from each variable in 2021 and 2019. Cluster 1 has the characteristics of provinces affected by COVID-19 very high in terms of changes in poverty. Cluster 1 has the characteristics of the highest average increase in Headcount Index (X_1), Poverty Gap Index (X_5), and Poverty Severity Index (X_6) compared to other clusters. Based on further investigation, it is known that 13 of the 16 provinces included in the cluster 1 are provinces that are still experiencing a contraction in the first quarter of 2021, which is marked by negative growth of the Gross Regional Domestic Product in the first quarter of 2021 (Year on Year/YoY). Bali, for example, experienced an economic contraction of up to 9.85% in January-March 2021. The economy of Central Java and West Java still experienced a decline of 0.87% and 0.83%, respectively. East Java's economy recorded a decline of 0.44%. With a sluggish economy, poverty will increase. Income which determines the purchasing power and price of consumer goods is an essential factor in determining the number and position of poverty. In addition, most of the provincial members in cluster 1 have a leading sector in the tourism sector. The COVID-19 pandemic has undoubtedly harmed the tourism industry and supporting sectors throughout Indonesia [17], so that the business field is directly affected by the tourism sector, which has an impact on poverty.

Cluster 2 consists of 17 provinces that are included in the moderate impact of COVID-19 on poverty. Cluster 2 has the characteristics of the highest average increase in Poverty line (X_2), Food Poverty Line (X_3), and Non-Food Poverty Line (X_4). The poor have an average monthly per capita expenditure below the poverty line, which is a line that shows the price value of basic needs. In 2021, including during the pandemic, the prices of consumer goods will not increase much. When viewed from the variables X1, X5, and X6, cluster 2 has an average increase that is not too high compared to cluster 1.

Cluster 3 has one member, namely Papua. Cluster 3 is the cluster that has a low COVID-19 impact on poverty. Based on Table 5, Cluster 3 shows a decrease in Headcount Index (X_1), Poverty Gap Index (X_5), and Poverty Severity Index (X_6) from 2019 to 2021. The COVID-19 pandemic has a negligible impact on cluster 3. Based on data from the Central Bureau of Statistics, in the first quarter of 2021, the Regional Domestic Product Growth Gross (GRDP) of Papua province grew by 14.28% (Year on Year), the highest in Indonesia. The GRDP expanded faster than the National economy, which shrank by -0.74% (YoY). In comparison to the fourth quarter of 2020, this growth rose by 6.92 percent (YoY). In the first quarter of 2021, Papua's economy grew mostly due to better performance in the mining sector, which coincided with an improvement in underground mining productivity, while the non-mining sector contracted by -3.76 percent (YoY). Household spending fell by -4.72 percent in terms of dollars spent (YoY). The decline in the non-mining sector was mostly caused by the spread of the COVID-19 virus, which harmed Papua's economic activity.

Based on the results of this study, the recommendation for the government is to pay more attention to the provinces in the first cluster. Because cluster 1 has characteristics that are prone to be affected by COVID-19 in terms of poverty and economy. In addition, economic recovery can be carried out through the tourism sector and MSMEs. Because of cluster 1, many provinces have excellent potential in the field of tourism. So that the priority sector in the tourism sector can boost the economy and reduce poverty. The provincial government can also implement several strategies, namely promoting the job market, doubling the benefits of cash transfers, maximizing the impact of fiscal stimulus, and encouraging green financing innovation.

CONCLUSIONS

In this study, we revealed that COVID-19 pandemic has the significant impact on poverty in Indonesia. The best method that produces the smallest Davies-Bouldin Index is Hybrid SOM-Kmeans using 3 clusters. The division of 3 clusters consists of high, moderate, and low impact of COVID-19 in poverty. Cluster 1 is a cluster with a high impact on poverty, which is a province with the primary sector in tourism. Because of the sluggish tourism, the purchasing power of the community is more diminutive and increases poverty. The cluster with a low impact is cluster 3, namely Papua, with the characteristics of the primary sector in the mining sector. Additional research into the effect of COVID-19 on each province is required. The author's recommendations, particularly for policymakers, are expected to enable the development of programs that are tailored to the characteristics of each cluster based on the analysis's findings.

REFERENCES

[1] Suryahadi A, Izzati R and Suryadarma D 2020 Estimating the Impact of Covid-19 on Poverty in Indonesia *Bull. of Ind. Econ. Studies* **56** 2, pp 175-192

[2] Tarigan H, Sinaga J H and Rachmawati R R 2020 Dampak Pandemi Covid-19: Perspektif Adaptasi dan Resiliensi Sosial Ekonomi Pertanian (Jakarta: IA ARD)

[3] Yusuf A A 2020 Poverty and distributional impact of Covid-19 crisis in Indonesia SDGs Center (Bandung: Universitas Padjadjaran)

[4] Johnson R and Wichern D 2014 Applied Multivariate Statistical Analysis (Sixth Edition) (Edinburg: Pearson)

[5] Han J, Kamber M and Pei J 2012 *Data Mining: Concepts and Techniques* (USA: Morgan Kaufmann)

[6] Ferezagia D V 2018 Analisis tingkat kemiskinan di Indonesia *Jurnal Sosial Humaniora Terapan* **1** 1 pp 1-6

[7] Sano A V D and Nindito H 2016 Application of K-Means Algorithm for Cluster Analysis on Poverty of Provinces in Indonesia *ComTech* **7** 2 pp 141-150

[8] Bahauddin A, Fatmawati A and Sari F P 2021 Analisis clustering provinsi di indonesia berdasarkan tingkat kemiskinan menggunakan algoritma k-means *Jurnal Manajemen informatika dan Sistem Informasi* **4** 1 pp 1-8

[9] Kohonen T 2001 *Self-Organizing Maps* (Berlin: Springer Series in Information Sciences)

[10] Kaski S and Kohonen T 1995 Exploratory data analysis by the Self-Organizing Map: structures of welfare and poverty in the world *the Third International Conference on Neural Networks in the Capital Markets* (London)

[11] Wilcoxon F 1945 Individual Comparisons by Ranking Methods *Biometrics Bulletin* **1** 6 pp 80-83
[12] Badran F, Yacoub M and Thiria S 2005 Self-organizing maps and unsupervised classification *Neural networks* pp 379-442
[13] Asan U and Ercan S 2012 An introduction to self-organizing maps *Computational Intelligence Systems in Industrial Engineering: with Recent Theory and Applications* (Paris: Atlantis Press) pp 295-315.
[14] Bates A and Kalita J 2016 Counting clusters in twitter posts Proceedings of the 2nd International Conference on Information Technology for Competitive Strategies
[15] Starczewski A 2017 A new validity index for crisp clusters *Pattern Analysis and Applications* **20** pp 687–700
[16] Wehrens R and Buydens L M C 2007 Self- and Super-organizing Maps in R: The kohonen Package *J. of Stat. Software* **21** 5 pp 1-19
[17] Pramana S, Paramartha D Y, Ermawan G Y, Deli N F and Srimulyani W 2021 Impact of COVID-19 pandemic on tourism in Indonesia *Current Issues in Tourism* **24**