Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

Analyzing COVID-19's Educational Impact in Indonesia: K-Means and Self-Organizing Map Approach

Ika Nur Laily Fitriana[1,*], Emeylia Safitri[2], Ria Faulina[3], Nuramaliyah[4], Fonda Leviany[5]

[1,2,3,4] Faculty of Science and Technology, Statistics Study Program, Universitas Terbuka, Tangerang Selatan, Indonesia
[5] Faculty of Science and Technology, Data Science Study Program, Universitas Terbuka, Tangerang Selatan, Indonesia
Email: [1,*]ika.nur@ecampus.ut.ac.id, [2]emeylia.safitri@ecampus.ut.ac.id, [3]ria.faulina@ecampus.ut.ac.id, [4]nuramaliyah@ecampus.ut.ac.id, [5]fonda.leviany@ecampus.ut.ac.id
Correspondence Author Email: ika.nur@ecampus.ut.ac.id

Abstract- The COVID-19 pandemic has affected the education sector. This research aimed to investigate the impact of COVID-19 on the education sector in Indonesia, especially on school participation indicators, using cluster analysis. We used fifteen factors related to the involvement indicators of students in elementary, junior secondary, and senior secondary education. The comparison of factors between 2019 and 2020 related to the effects of COVID-19, which began to proliferate in Indonesia in March 2020. Consequently, comparing those periods yields insights into the timeframe before and after the spread of COVID-19. To assess the pandemic's influence on the education sector, we performed an inferential statistical analysis using a nonparametric location test to identify significant changes between variables in 2019 and 2020. Subsequently, we performed cluster analysis using K-Means and Self-Organizing Map (SOM) approaches. The optimal cluster obtained for K-Means and SOM is three clusters. The results indicate that SOM and K-Means exhibit similar performances. Changes in cluster members in 2019 and 2020 indicate an enormous impact due to COVID-19. Cluster 3, which consists of DKI Jakarta, West Java, Central Java, East Java, and North Sumatra, is most affected by the pandemic from the educational sector.

Keywords: Cluster; Education; K-Means; Pandemic; SOM

1. INTRODUCTION

In early 2020, Indonesia recorded its first COVID-19 cases and rapidly entered a prolonged public-health emergency that strained health and social systems nationwide. By 21 July 2021, official reports tallied 2,983,830 confirmed cases and 77,583 deaths across all 34 provinces, underscoring the scale and geographic spread of the crisis. Education was among the sectors most profoundly affected: emergency policy (Ministerial Circular No. 4/2020) mandated learning-from-home arrangements and suspended routine face-to-face instruction, compelling schools to pivot to distance modalities [1]. Thus, the entire teaching and learning process takes place via distance education mechanisms such as online learning and educational television with limited connectivity, yet uptake and access varied markedly across regions and socioeconomic groups. Documented barriers included the high cost and unstable quality of internet access, scarcity of devices, and uneven digital skills, especially outside major urban centers [2].

The magnitude of disruption was unprecedented for Indonesia's approximately 68 million school-age learners, many of whom experienced long school closures and substantial instructional time loss [3]. Emerging empirical work indicates that these disruptions translated into measurable learning deficits and heightened risks to educational participation, especially for older adolescents and students from lower-income households [4]. Household shocks and constrained resources during the economic contraction further threatened school participation and re-enrollment, raising concerns about dropout and disengagement [5]. United Nations Children's Fund (UNICEF) monitoring observed that roughly 1% of school-age respondents had dropped out during the pandemic period, with economic reasons frequently cited; other national assessments place the figure in a similar order of magnitude. At the system level, Indonesia's PISA 2022 results, which reflect learning accumulated through the pandemic, showed declines across mathematics, reading, and science relative to 2018, mirroring global patterns but highlighting persistent equity gaps and a relatively small proportion of students reaching baseline proficiency [6]. Together, these findings motivate a closer, quantitative examination of how the pandemic altered school participation across Indonesia's diverse provinces and districts.

Beyond narrative and policy analyses, clustering methods can add value by revealing structure in multi-indicator education data, identifying geographic profiles and temporal shifts that may be obscured in aggregate statistics. This study relies on cluster analysis to assess the impact of COVID-19 on educational sectors, focusing on indicators of school participation. Research has been conducted to evaluate the impact of COVID-19 on the education sector. Abidin and Tobibatussa'adah [3] analyzed the impact of the COVID-19 pandemic on educational and judicial practices in Indonesia. Rulandari [4] conducted a study on the effects of COVID-19 in Indonesia employing a descriptive-qualitative methodology. Abidah et al [1] conducted a conceptual analysis to develop a position paper addressing the impact of COVID-19 on Indonesian education and its relation to the philosophy of "merdeka belajar." The majority of this research employed qualitative methodologies, encompassing descriptive and policy analysis. More recent studies have further documented learning disruptions and participation challenges during the pandemic. For example, Ayuningtyas and Arsana [7] identified significant barriers to distance learning, including physical, psychological, facilities and learning materials. Azzahra [8] highlighted widening educational gaps associated with unequal access to technology. Recent international evidence suggests that the effects of COVID-19 on school participation persist beyond the acute phase of the pandemic, with several countries facing delayed re-entry and uneven recovery across regions [9]-[10]. The World Bank [11]further

Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

underscores that learning poverty and participation gaps remain pronounced in middle-income countries, reinforcing the need for targeted, data-driven recovery strategies. These findings underscore the importance of examining subnational participation patterns using multi-indicator approaches.
No previous studies have quantitatively examined the impact of COVID-19 on the education sector in Indonesia through cluster analysis. Consequently, the researcher wants to examine several indicators of school involvement before and during the pandemic using clustering methods, notably K-means clustering and Self-Organizing Maps (SOM) approach. K-means clustering provides a simple, scalable partitioning of observations based on similarity, while Self-Organizing Maps (SOM) jointly perform topology-preserving projection and clustering[12], facilitating interpretation of high-dimensional patterns on a two-dimensional map. Although clustering has been applied to COVID-19 epidemiological risks in Indonesia and to education-related groupings in selected settings, its use to characterize pre- versus mid-pandemic school-participation indicators at scale remains limited. Building on these methodological strengths, the present study employs K-means and SOM to (i) profile Indonesian regions by school-participation indicators before and during COVID-19, (ii) visualize changes in participation patterns during the disruption, and (iii) generate actionable segmentation that can inform targeted recovery policies and resource allocation.

2. RESEARCH METHODOLOGY

2.1 Step of Analysis

This study explores the impact of COVID-19 in the education sector. This research uses quantitative data analysis research with a statistical approach. The general analysis steps are depicted in the flow chart in Figure 1.

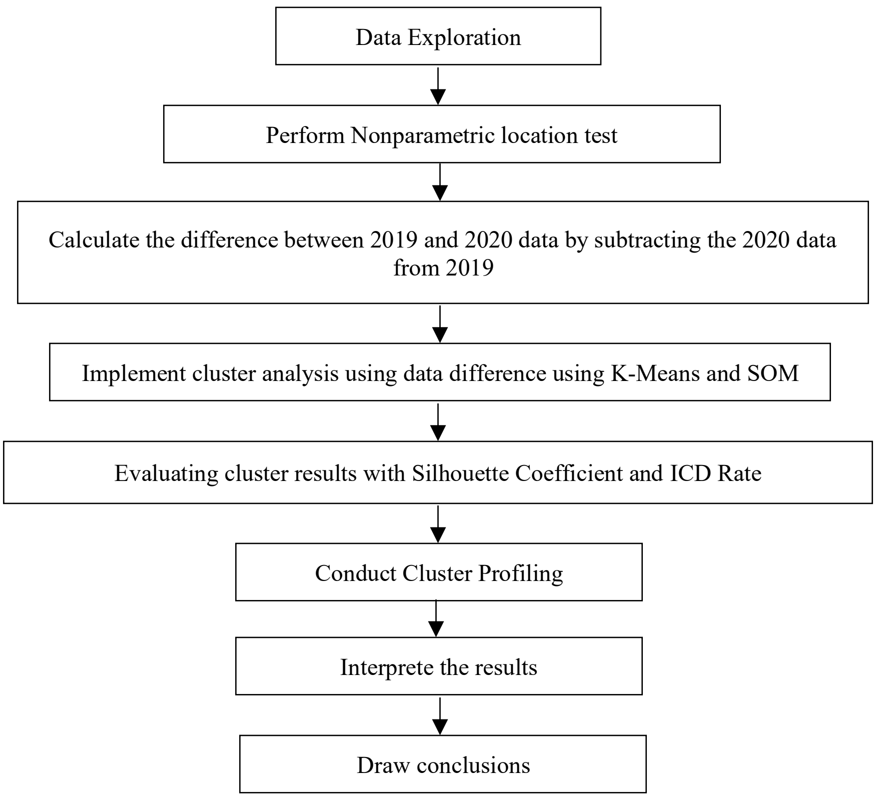

Figure 1. Flow Chart

As an application of the variables used, the 2019 data will be compared with the data in 2020. Comparison of variables in 2019 and 2020 is related to the impact of COVID-19, which has been spreading through Indonesia on March 2020. Thus, the comparison of those times provides information before and after the spread of COVID-19. 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 2020). Then, clusters were carried out using K-Means and SOM methods. Furthermore, the cluster's performance was evaluated using the silhouette coefficient. The explanation of each method used is as follows.

2.1.1 Wilcoxon Signed-Ranked Test

When the data do not satisfy the assumption of (multivariate) normality, hypothesis testing on differences in location parameters can be conducted using a nonparametric alternative. Accordingly, this study applies the Wilcoxon

Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

signed-rank test, a distribution-free procedure designed to evaluate whether the median of paired differences differs from zero for paired observations measured on at least an ordinal (and commonly interval) scale. The Wilcoxon signed-rank test is widely used as a robust alternative to the paired t-test when the normality assumption is violated [13]. In this study, the null hypothesis is median of differences is equal to zero and alternative hypothesis is median of differences is not equal to zero If probability (p-value) < 0.05, then null hypothesis is rejected. If probability (p-value) > 0.05 then null hypothesis is not rejected.

2.1.2 K-Means Clustering

K-means clustering method is a non-hierarchical (partition-based) technique that partitions a set of observations into a predefined number of clusters (K) based on similarity among their characteristics [13]. The objective of this method is to maximize homogeneity within clusters while ensuring heterogeneity between clusters. The clustering procedure begins with the specification of the number of clusters K, followed by the initialization of cluster centroids. Each observation is then assigned to the cluster whose centroid is closest, typically measured using Euclidean distance. After the initial assignment, cluster centroids are recalculated as the mean of all observations within each cluster. This process of assigning observations and updating centroids is iteratively repeated until convergence is achieved, indicated by stable cluster memberships or minimal changes in centroid positions. The final result is a partition of the data into K clusters that best represent the underlying structure of the dataset [14].

2.1.2 Self-Organizing Map (SOM)

The Self-Organizing Map (SOM) is an unsupervised artificial neural network designed to group and visualize high-dimensional data by preserving topological relationships among observations. Through competitive learning, SOM represents clusters as nodes on a low-dimensional map, enabling intuitive comparison of patterns and similarities across groups [15], [16]. The SOM training procedure can be summarized as follows. First, the dimensionality and size of the output map are specified, and initial weight vectors are assigned to each node. At each iteration t , an input vector x(t) is randomly sampled from the training dataset. The similarity between the input vector and all node weight vectors is then evaluated, typically using the Euclidean distance, and the best matching unit (BMU) is identified as the node with the minimum distance:

$$c(t) = argmin\{\|x(t) - w_i(t)\|\}, i = 1,2, \dots , n \quad (1)$$

Subsequently, the weight vectors of the BMU and its neighboring nodes are updated according to their proximity to the BMU, using a learning rate $\alpha(t)$and a neighborhood function $h_{ci}(t)$, where $h_{ci}(t) = 1$for the winning node:

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

After each iteration, the learning rate and neighborhood radius are gradually reduced to allow the model to transition from capturing global structures to refining local patterns. The process continues until a convergence criterion is met, such as negligible changes in weight vectors or the attainment of a predefined maximum number of iterations.

2.1.4 Cluster Evaluation (Silhouette Coefficient and ICD Rate)

Silhouette Coefficient is a method used to evaluate the results of clustering by checking how well the resulting clusters [17]. Equation (3) is the formula for calculate silhouette coefficient.

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (3)$$

Where $a(i)$ is average distance from the i-th observation with all observations in the same cluster. $b(i)$ is the smallest value of average distance from the i-th observation with all observations in different clusters. A value for the silhouette coefficient that is close to one indicates that the cluster is performing well [18]. Evaluation of the best cluster performance results can be determined using the internal cluster dispersion rate or commonly known as the ICD rate. In this study, the ICD rate is defined using the ratio of between-cluster variability to total variability. ICD rate calculated using equation (4) is used to measure the level of dispersion between clusters formed. The smaller the ICD rate, the better the grouping results.

$$ICD\ Rate = 1 - \frac{SSB}{\mathrm{SST}} \quad (4)$$

where SSB is the between-cluster sum of squares and SST is the total sum of squares. A smaller ICD rate indicates that a larger proportion of the total variability is explained by between-cluster differences, implying more distinct cluster separation and, consequently, better clustering structure.

Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

2.2 Research Variables

We use secondary data from two official sources for 2019 (pre-pandemic) and 2020 (first pandemic year in Indonesia): (i) national education statistics compiled by BPS-Statistics Indonesia (derived largely from SUSENAS and administrative registers), and (ii) administrative education statistics published by the Ministry of Education and Culture (Kemendikbud). These publications report school participation and dropout indicators at the provincial level, which comprises 34 provinces. This research enumerates the variables in Table 1.

Table 1. Research variables

Symbols	Variables
X_1	Number of primary-school pupils
X_2	Number of junior secondary-school pupils
X_3	Number of senior secondary-school pupils
X_4	Number of dropouts – primary
X_5	Number of dropouts – junior secondary
X_6	Number of dropouts – senior secondary
X_7	GER – primary
X_8	GER – junior secondary
X_9	GER – senior secondary
X_{10}	NER – primary
X_{11}	NER – junior secondary
X_{12}	NER – senior secondary
X_{13}	Transition rate to junior secondary
X_{13}	Transition rate to senior secondary
X_{15}	% of youth aged 10–24 whose main activity is household care in the last week (proxy for at-risk disengagement)

3. RESULT AND DISCUSSION

3.1 Data Exploration

Figure 2 provides a descriptive summary of selected education indicators in Indonesia. The figure suggests that the total number of pupils across all education levels remained broadly stable between 2019 and 2020. In contrast, several indicators exhibit declines over the same period, including the number of dropouts at all levels, the Gross Enrolment Ratio for primary school (GER PS), the Net Enrolment Ratios for primary, junior secondary, and senior secondary education (NER PS, NER JSS, and NER SSS), as well as the transition rates to junior secondary and senior secondary schooling (Transition JSS and Transition SSS). Notably, despite the observed reduction in dropout counts, participation-related indicators also declined, indicating that lower dropout figures did not necessarily correspond to improved engagement in schooling. Meanwhile, the householding indicator increased from 2019 to 2020, implying that during the pandemic a larger share of youth reported household-related activities as their primary activity in the preceding week.

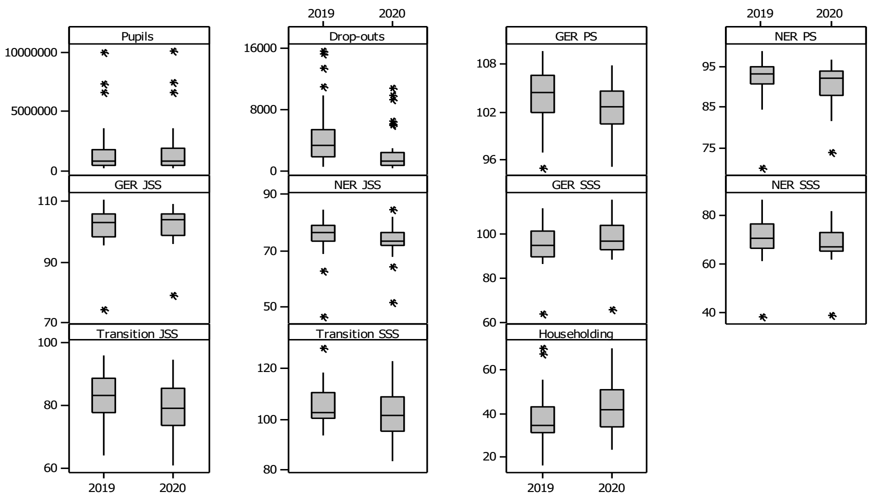

Figure 2. Boxplot of education indicators in 2019 and 2020.

Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

3.2 Difference Test Analysis

In this section, we provide multivariate difference test analysis for all variables in 2019 and 2020 using hypothesis testing. Before doing difference tests, we perform multivariate normality tests. For all variables in 2019 and 2020, the multivariate normality test shows that the data does not follow the multivariate normal distribution, so that we use nonparametric method, namely Wilcoxon Signed-Ranked Test. The nonparametric location test was used in this study to determine the significant differences between variables in 2019 and 2020. We found that the median of differences is not equal to zero using the multivariate nonparametric location test, with a statistic of 5.0862 and a p-value <0.001. This implies that the data for 2019 and 2020 are multivariate significantly different. Table 2 shows the results of univariate nonparametric location tests. All variables are significantly different except for the Gross Enrolment Ratio (GER) for senior secondary schools (X_9). All variables except X_9 have p-value less than significant level (0.05). Estimated values with a positive sign indicate a decrease between 2019 and 2020.

Table 2. Result of nonparametric location test.

Variables	estimates	statistics	p-value	Variables	estimates	statistics	p-value
X_1	436.50	3.907	0	X_9	1.44	1.735	0.079
X_2	2.90	5.086	0	X_{10}	-2.26	-4.573	0
X_3	624.50	5.052	0	X_{11}	2.87	3.069	0.002
X_4	3.69	4.539	0	X_{12}	2442.5	2.368	0.017[a]
X_5	754	4.385	0	X_{13}	-4186	-3.189	0.001
X_6	1.72	3.530	0	X_{14}	-6969	-4.932	0
X_7	1.37	2.607	0.008	X_{15}	-4.63	-4.727	0
X_8	-0.59	-4.505	0				

[a]Non-significant difference

3.3 Cluster Analysis

In this section, we provide cluster analysis using K-Means and SOM methods. Clustering is done based on the variables that have been determined. First, the optimal number of clusters is selected based on the highest silhouette coefficient for each method. According to Table 3, the highest silhouette coefficient using the k-means method in 2019 data is 0.450256 with the optimum number of clusters being 3 clusters. Similarly, for 2020 data using the K-means method, the highest silhouette coefficient obtained is 0.432422, with the optimum number of clusters being 3 clusters.

Table 3. Silhouette coefficient of K-Means.

Number of Clusters	2019	2020
2	0.426764	0.41933
3	0.450256[a]	0.432422[a]
4	0.236649	0.177788
5	0.205733	0.200895

[a]The highest value

The number of clusters in SOM methods is equal to the multiplication of grid dimension. According to Table 4, the highest silhouette coefficient in 2019 data is 0.450 for a 3 x 1 grid size, indicating that 3 clusters are the optimal number. In the 2020 data, the optimal cluster is also 3 clusters, as it has the highest silhouette coefficient.

Table 4. Silhouette coefficient of SOM.

Grid Size	2019	2020	Grid Size	2019	2020	Grid Size	2019	2020
2 x 1	0.427	0.419	1 x 3	0.450[a]	0.432[a]	1 x 4	0.175	0.175
1 x 2	0.427	0.419	4 x 1	0.227	0.175	5 x 1	0.177	0.18
3 x 1	0.450[a]	0.432[a]	2 x 2	0.366	0.195	1 x 5	0.198	0.146

[a]The highest value

After selecting the number of clusters, we used the ICD Rate criteria to compare the performance of the best method to K-mean and SOM. The best method is the method that has the smallest ICD Rate, which indicates that the differences between groups are slight. Table 5 summarizes the results of the ICD rate calculation. The ICD rate is the same for both K-Means and SOM clustering, implying that the two methods perform similarly in grouping 34 provinces based on education indicators in 2019 and 2020. As a result, the member of the cluster is equal when K-Means and SOM are used.

Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

Table 5. Method evaluation using ICD Rate.

Methods	2019		2020	
	Number of Clusters	ICD Rate	Number of Clusters	ICD Rate
K-Means	3	0.2081093	3	0.3315759
SOM	3	0.2081093	3	0.3315759

Details of the provincial members from each cluster are shown in Table 6. The provincial map is shown in Figure 2. Cluster 1 has only one member, namely Papua. We can see that there were changes in cluster members in 2019 and 2020. In 2019, DKI Jakarta Province was included in cluster 2, but in 2020, DKI Jakarta was included in cluster 3.

Table 6. Provincial members in each cluster.

Year	Cluster	Cluster Members (Provinces)	Number of Members
2019	1	Papua	1
	2	Aceh, West Sumatra, Riau, Jambi, South Sumatra, Bengkulu, Lampung, Bangka Belitung Islands, Riau Islands, DKI Jakarta, DI Yogyakarta, Banten, Bali, West Nusa Tenggara, East Nusa Tenggara, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, Central Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo, West Sulawesi, Maluku, North Maluku, West Papua	29
	3	North Sumatra, West Java, Central Java, East Java	4
2020	1	Papua	1
	2	Aceh, West Sumatra, Riau, Jambi, South Sumatra, Bengkulu, Lampung, Bangka Belitung Islands, Riau Islands, DI Yogyakarta, Banten, Bali, West Nusa Tenggara, East Nusa Tenggara, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, Central Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo, West Sulawesi, Maluku, North Maluku, West Papua	28
	3	North Sumatra, DKI Jakarta, West Java, Central Java, East Java	5

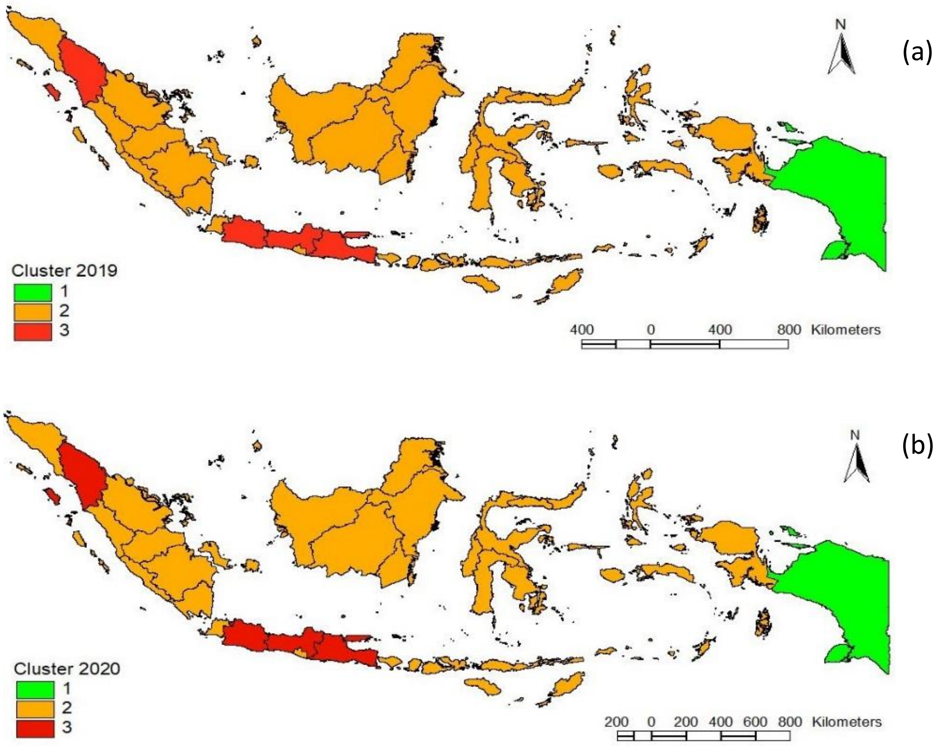

Figure 2. Map of Provinces using Cluster Analysis in (a) 2019 and (b) 2020.

Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

The identification of characteristics in each cluster is shown in Table 7. In 2019 and 2020, Papua became the only member in cluster 1 because Papua has the most different characteristics from other provinces in Indonesia. Based on the Central Bureau of Statistics data, the Papua Province Human Development Index (IPM) is the lowest in Indonesia, 60.84 in 2019 and 60.44 in 2020. Sukiastini [19] revealed that before the COVID-19 pandemic, education in Papua was low and a high disparity compared to other provinces due to limited access, infrastructure, facilities, and teaching staff. Based on Paling and Sitorus's research [20], the effectiveness of online learning in Papua during the COVID-19 pandemic was ineffective in improving student learning outcomes. The most significant impact of the COVID-19 in cluster 1 (Papua) is seen in the transition rate to junior secondary, which in 2020 saw a contraction of 19%. In addition, the COVID-19 pandemic also has an impact on an increase in students doing household activities in cluster 1 by an increase of this variable is the largest compared to other clusters. It is in line with Sukiastani's research [19], which states that due to the COVID-19 pandemic, many students in Papua have replaced their parent's role. However, the good news is that there is no decrease in students at all levels in cluster 1. Furthermore, the decrease of drop-out students at the elementary school is the highest among other clusters. There is a reduction in the number of children aged 7-12 years who are no longer in school or have not completed their education at a primary school level.

Based on the analysis results, there was a change in members in cluster 2 in 2019, which originally numbered 29, to 28 members. This change is due to changes in the characteristics of the DKI Jakarta province due to the COVID-19 pandemic. Cluster 2 is categorized as a middle cluster due to the impact of COVID-19 on student numbers at all levels is not relatively high. Meanwhile, in the DKI Jakarta province, the impact of COVID-19 is quite significant, so that DKI Jakarta became Cluster 3 in 2020.

Cluster 3 has the characteristics of the average number of students at the elementary to the senior high school level is high, because the provinces included in cluster 3 are provinces with a crowded population. Cluster 3 is the cluster most affected by COVID-19 in terms of education, especially in student participation, because it has the highest contraction in the number of students at all levels from the previous year. According to Table 7, the number of primary school students decreased by 15%, or 528,605 students, from the previous year. The number of junior secondary school pupils decreased by 13% or 229,962. Senior secondary school registration dropped by 12% or 184,246 students. The member of cluster 3 are DKI Jakarta, West Java, Central Java, East Java, and North Sumatra and. These provinces are the highest COVID-19 cases in Indonesia.

In general, from the three clusters, there was an increase in the number of students taking care of the household. The activities of taking care of the household include taking care of help or take care of the household, such as cooking, washing, and cleaning the house. Apart from studying, other student activities can be in the form of helping with household chores. This study revealed that there is impact of pandemic on education sector, in line with research from Upoalkpajor and Upoalkpajor [21].

Table 7. Percentage of changing of average some variables in each cluster.

Variables	Cluster 1	Cluster 2	Cluster 3
Number of primary school’s pupils	0%	-4%	-15%
Number of junior secondary school’s pupils	0%	-2%	-13%
Number of senior secondary school’s pupils	0%	-1%	-12%
Number of drop-outs in primary school	-72%	-39%	-8%
Transition rate to junior secondary school	-19%	-5%	0%
Percentage of pupils who take care of household	28%	14%	2%

4. CONCLUSION

In this study, we found the impact of COVID-19 on the education sector, especially in student participation. Based on the variables determined in this study, using a nonparametric location test shows significant multivariate differences before and after COVID-19. Cluster analysis using K-Means and SOM methods results in the optimum number of clusters being 3 clusters. The result shows that SOM and K-Means have the same performance. There are changes in cluster members in 2019 and 2020 which indicate an enormous impact due to COVID-19. Cluster 3, which consists of DKI Jakarta, West Java, Central Java, East Java, and North Sumatra, is most affected by the pandemic. Further research in investigating the effect of COVID-19 for each province is necessary to be conducted.

REFERENCES

[1] A. Abidah, H. N. Hidaayatullaah, R. M. Simamora, D. Fehabutar, L. Mutakinati, and N. Suprapto, “The Impact of Covid-19 to Indonesian Education and Its Relation to the Philosophy of ‘Merdeka Belajar,’” *Studies in*

Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

Philosophy of Science and Education (SiPoSE), vol. 1, no. 1, pp. 38–49, 2020, [Online]. Available: http://scie-journal.com/index.php/SiPoSE

[2] J. E. Wallace, "The Impact of Remote Learning on Student Engagement and Academic Performance during the COVID-19 Pandemic," *Journal of Education and Teaching Methods*, vol. 2, no. 2, pp. 20–32, 2023.

[3] Z. Abidin and Tobibatussa'adah, "THE IMPACT OF COVID-19 PANDEMIC ON EDUCATION AND JUDICIAL PRACTICE IN INDONESIA," *RI'AYAH*, vol. 5, no. 2, pp. 122–130, 2020, Accessed: Jan. 31, 2026. [Online]. Available: https://e-journal.metrouniv.ac.id/riayah/article/view/2794/1886

[4] N. Rulandari, "The Impact of the Covid-19 Pandemic on the World of Education in Indonesia," *IJSS Ilomata International Journal of Social Science*, vol. 1, no. 4, pp. 242–250, 2020, [Online]. Available: https://www.ilomata.org/index.php/ijss

[5] R. Afkar and N. Yarrow, *Rewrite the future: How Indonesia's education system can overcome the losses from the COVID-19 pandemic and raise learning outcomes for all*. Washington DC: World Bank, 2021. Accessed: Jan. 31, 2026. [Online]. Available: https://documents1.worldbank.org/curated/en/589551630680730676/pdf/Rewrite-the-Future-How-Indonesias-Education-System-can-Overcome-the-Losses-From-the-COVID-19-Pandemic-and-Raise-Learning-Outcomes-for-All.pdf

[6] OECD, "PISA 2022 Results Factsheets Indonesia," 2023. Accessed: Jan. 31, 2026. [Online]. Available: https://oecdch.art/a40de1dbaf/C108.

[7] V. Ayuningtyas and I. M. Arsana, "Hambatan Pembelajaran Jarak Jauh di Masa Pandemi Covid-19," *Jurnal Pendidikan Teknik Mesin*, vol. 11, no. 2, pp. 186–194, 2022.

[8] N. F. Azzahra, "Addressing Distance Learning Barriers in Indonesia Amid the Covid-19 Pandemic," Jakarta, 2020. Accessed: Jan. 31, 2026. [Online]. Available: https://hdl.handle.net/10419/249436

[9] OECD, "Education at a Glance 2024: OECD Indicators," *Education at a Glance*, vol. 2024, Sep. 2024, doi: 10.1787/C00CAD36-EN.

[10] UNESCO, "Education: From disruption to recovery | UNESCO." Accessed: Jan. 31, 2026. [Online]. Available: https://www.unesco.org/en/covid-19/education-disruption-recovery

[11] World Bank, "The State of Global Learning Poverty: 2022 Update," World Bank. Accessed: Jan. 31, 2026. [Online]. Available: https://www.worldbank.org/en/topic/education/publication/state-of-global-learning-poverty

[12] A. Karim, S. Esabella, K. Kusmanto, M. Hidayatullah, and S. Suryadi, "Penerapan Data Mining Untuk Pengelompokan Terhadap Kualitas Kinerja Karyawan Dengan Menggunakan Algoritma K-Medoids Clustering," *Jurnal Media Informatika Budidarma*, vol. 8, no. 2, p. 1001, 2024, doi: 10.30865/mib.v8i2.7445.

[13] J. D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference*, 6th ed. New York: Chapman and Hall/CRC, 2020. doi: 10.1201/9781315110479.

[14] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci. (N. Y).*, vol. 622, pp. 178–210, Apr. 2023, doi: 10.1016/j.ins.2022.11.139.

[15] S. Licen, A. Astel, and S. Tsakovski, "Self-organizing map algorithm for assessing spatial and temporal patterns of pollutants in environmental compartments: A review," *Science of The Total Environment*, vol. 878, p. 163084, Jun. 2023, doi: 10.1016/J.SCITOTENV.2023.163084.

[16] T. Kohonen, *Self-Organizing Maps*, vol. 30. in Springer Series in Information Sciences, vol. 30. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001. doi: 10.1007/978-3-642-56927-2.

[17] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. USA: Morgan Kaufmann, 2012. doi: 10.1016/B978-0-12-381479-1.00001-0.

[18] D. Xu and Y. Tian, "A Comprehensive Survey of Clustering Algorithms," *Annals of Data Science 2015 2:2*, vol. 2, no. 2, pp. 165–193, Aug. 2015, doi: 10.1007/S40745-015-0040-1.

Bulletin of Information Technology (BIT)
Vol 7, No 1, Maret 2026, Hal. 31 - 39
ISSN 2722-0524 (media online)
DOI 10.47065/bit.v5i2.2581
https://journal.fkpt.org/index.php/BIT

[19] I. G. A. N. K. Sukiastini, “Dunia Pendidikan di Wilayah Pedalaman Papua Sebelum dan Setelah Terdampak Covid-19,” *Syntax Idea*, vol. 2, no. 8, pp. 381–388, 2020, Accessed: Feb. 01, 2026. [Online]. Available: https://jurnal.syntax-idea.co.id/index.php/syntax-idea/article/view/497/392

[20] S. Paling and M. Sitorus, “Efektifitas Pembelajaran Daring Pada Masa Pandemi Covid-19 di Papua,” *Jurnal Education and development*, vol. 9, pp. 64–71, May 2021.

[21] J.-L. N. Upoalkpajor and C. B. Upoalkpajor, “The Impact of COVID-19 on Education in Ghana,” *Asian Journal of Education and Social Studies*, pp. 23–33, Jun. 2020, doi: 10.9734/AJESS/2020/V9I130238.