ANALYSIS OF REGIONAL CLUSTERS IN INDONESIA BASED ON FACTORS CAUSING STUNTING USING SELF ORGANISING MAP ALGORITHM

Karyati* and Larrachita Rizka Bellapurani

Department of Mathematics Education, Mathematics and Natural Science Faculty of Universitas Negeri Yogyakarta, Yogyakarta Indonesia

Article Info

Article history:
Received July 24th, 2024
Revised January 22nd, 2025
Accepted February 23rd, 2025

****Corresponding Email:***
karyati@uny.ac.id

ABSTRACT

This study aimed to classify regions in Indonesia based on the causes of stunting which are included in specific intervention and sensitive intervention framework, so that related parties can use them to address the biggest causal factors in each region. This research used the Self Organizing Map (SOM) method. The variables used were the percentage of children aged less than 6 months who were exclusively breastfed, the percentage of children aged 12-23 months who received complete basic immunization, the percentage of ever-married women aged 15-49 years whose last birth was facilitated by health workers and assisted by health workers, the percentage of households that have access to proper drinking water sources, and the percentage of households that have access to proper sanitation. The results obtained 6 clusters with their respective characteristics. Cluster 1 consists of 27 districts/cities, cluster 2 consists of 59 districts/cities, cluster 3 consists of 23 districts/cities, cluster 4 consists of 264 districts/cities, cluster 5 consists of 103 districts/cities and cluster 6 consists of 38 districts/cities.

Keywords: Stunting, Clustering, Self Organizing Map

Introduction

Accelerating the reduction in stunting cases is one of the government's programs in the health sector. Stunting is a condition of growth failure in children under five years of age (toddlers) due to chronic malnutrition, especially during the first 1,000 days of life (HPK), resulting in toddlers being shorter than average for their age. Stunting is caused by many factors such as socioeconomic conditions, maternal nutrition during pregnancy, congenital diseases in infants, inadequate nutritional intake, and inadequate environmental conditions. This causes toddlers suffering from stunting to have difficulty achieving optimal physical and cognitive development in the future (Indonesian Ministry of Health, 2018). Stunting can occur in toddlers born to teenage mothers. It becomes severe when pregnant women do not consume enough nutrients, especially when they live in environments with poor sanitation. Ultimately, the harmful effects of stunting are that it hinders economic growth, increases poverty, and widens inequality.

Factors contributing to stunting include poor parenting practices, such as parents' lack of knowledge about health and nutrition before and during pregnancy. Children aged 0-6 months who are not exclusively breastfed. In addition, environmental factors are also one of the causes of stunting. These environmental factors include inadequate health facilities, lack of access to clean water and proper sanitation for households/families (Kominfo, 2019).

In this study, the author will cluster areas that are prioritized in efforts to address the factors causing stunting in order to reduce stunting rates in Indonesia. The indicators of stunting causes used are in line with the nutrition-sensitive and specific interventions established by the government, namely the percentage of children under 6 months of age who are exclusively breastfed, the percentage of children aged 12-23 months who receive complete basic immunizations, the percentage of married women aged 15-49 years whose last delivery was facilitated by a health facility and assisted by health personnel, the percentage of households with access to safe drinking water sources, and the percentage of households with access to proper sanitation. Cluster analysis aims to create clusters or groups of data that share similar characteristics within a cluster or group and have different characteristics from other clusters, taking into account the distance or measure of dissimilarity between data. The set of clusters resulting from cluster analysis can be referred to as clustering (Han *et al*., 2012). One method that can be used to group data is Self Organizing Map (SOM). SOM is included in unsupervised learning algorithms to produce input representations in two dimensions. SOM has an architecture consisting of two layers, namely input and output that are interconnected. Each neuron in the output represents a cluster of inputs (Windarto et al., 2020).

Research on stunting has been studied by several researchers, including research conducted by Maimunah *et al.* (2023). In this study, data clustering on the prevalence of stunting in Tegalrejo Village, Magelang Regency, was carried out with the aim of reducing the prevalence of stunting in the village. The method used was Agglomerative Hierarchical Clustering, which produced three clusters: high, medium, and low prevalence of stunting. The next study on stunting was conducted by Pohan *et al.* (2021), who used the K-Medoids algorithm to group stunted toddlers based on province in Indonesia. The purpose of this study was to group toddlers experiencing stunting into two clusters: high and low. Another study was conducted by the Central Statistics Agency (BPS) in the Central Statistics Agency Health Statistics Profile (2019), which discussed the grouping of priority districts/cities for stunting. The study was conducted using cluster analysis methods. The variable used in this study was secondary data from BPS in 2018. However, in the study conducted by BPS, the cluster results still showed classification errors, where there were areas that should have been prioritized but were not included in the priority list.

Previous research that used the SOM algorithm in the clustering process was conducted by Mardiyah (2022), who used SOM in grouping of COVID-19 spread areas in East Java province resulted in 3 clusters. Another study using the SOM algorithm was conducted by Firmansyah et al. (2019) to group areas based on social welfare. In this study, the Missing Value K-Nearest Neighbors (KNN) improvement was carried out to overcome the problem of a large amount of missing data in the dataset, and 2 clusters were obtained. Another study was conducted by Suwirmayanti (2020), who used SOM to group computer engineering students at STIKOM Bali based on their course preferences. The parameters used in this study were the previous semester's course grades, which were included in the prerequisites for selecting majors, and 2 clusters were produced. Another study was conducted by Kasih and Rizki (2019), who used SOM to cluster public complaints on Radio X in Kediri City. This study grouped public complaints related to government agencies that interact with the community, namely the Health Office, Education Office, Civil Registry Office, Public Works Office, and Tourism Office. Modeling used text mining techniques to pre-process data in the form of sentences, while SOM was used to represent the data and determine the appropriate agency group for public complaints.

This article reports on regions in Indonesia based on indicators of factors causing stunting using the Self-Organizing Map algorithm. Therefore, this thesis is titled Cluster Analysis of Regions in Indonesia Based on Factors Causing Stunting Using the Self-Organizing Map Algorithm.

Methods

Research Data

The data used in this study is secondary data obtained from the official website of the Central Statistics Agency (BPS), namely www.bps.go.id. The data consists of 514 districts/cities with 5 variables, including the percentage of children under 6 months of age who are exclusively breastfed, the percentage of children aged 12-23 months who have received complete basic immunizations, the percentage of married women aged 15-49 years whose last delivery was facilitated by a health facility and assisted by health personnel, the percentage of households with access to safe drinking water sources, and the percentage of households with access to proper sanitation, as described in Table 1.

Table 1. Data variable

No	Variable	Unit	Data Sample
1	Children under 6 months of age who are exclusively breastfed according to district/city (X_1)	Percent	The percentage of children under six months of age who are exclusively breastfed in Simelue district is 70%. Meanwhile, in Aceh Singkil district, the percentage is 82%
2	Children aged 12-23 months who have received complete basic immunizations according to district/city (X_2)	Percent	The percentage of children aged 12-23 months who received complete basic immunizations in Simelue district was 65%. Meanwhile, in Aceh Singkil district, it was 53%
3	Women aged 15-49 who have been married and whose last delivery was facilitated and assisted by health personnel, by district/city (X_3)	Percent	he percentage of women aged 15-49 who have ever been married and whose last delivery was facilitated and assisted by health workers in Simelue district is 59%. Meanwhile, in Aceh Singkil district, the percentage is 60%
4	Households with access to proper drinking water sources according to the district/city (X_4)	Percent	The percentage of households with access to proper drinking water sources in Simelue District is 77.16%. Meanwhile, in Aceh Singkil District, it is 62.27%.
5	Households with access to adequate sanitation according to district/city (X_5)	Percent	The percentage of households with access to proper sanitation in Simelue District is 65.47%. Meanwhile, in Aceh Singkil District, it is 73.54%.

Working Procedures

Data grouping in this study used the SOM algorithm, which was calculated using R Studio software. The research flow is described in Figure 1.

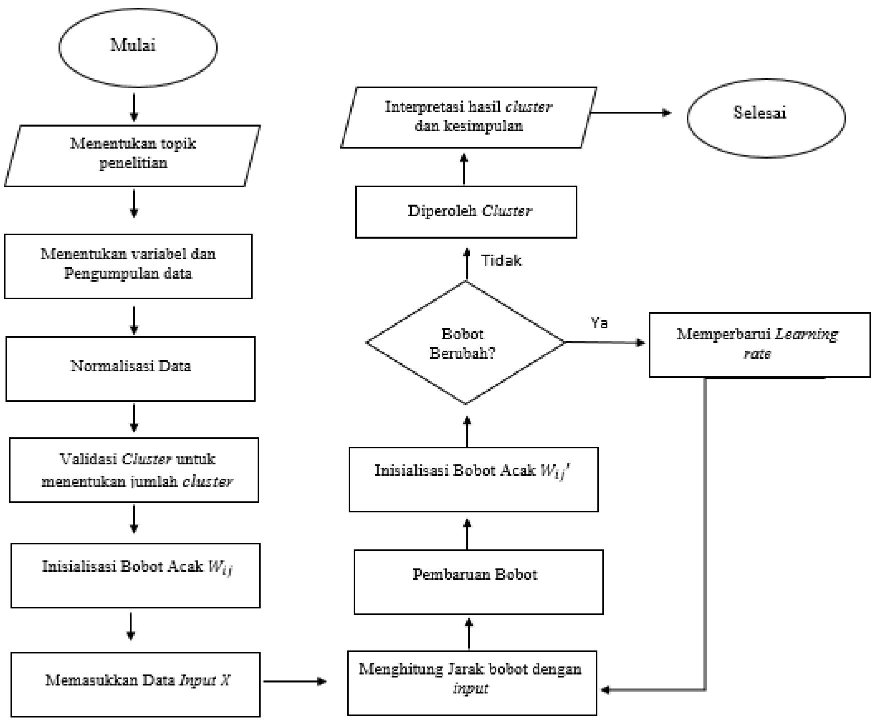

Figure 1. Research steps chart

Based on Figure 1, the research steps are as follows

1. Determine the research topic, determine the variables, and collect the data to be used.
2. Data Normalization

   Data normalization is performed to reduce the data range. In this study, data normalization is performed using the Z-Score Normalization or Standardization method. Standardization has a range between -3 and 3. The normalization process is assisted by RStudio software.
3. Validate the number of clusters

   The best number of clusters can be determined through cluster validation. There are several types of cluster validation, one of which is internal cluster validation. The types of internal cluster validation are as follows

   a. *Dunn* index

   The *Dunn* index has a value between 0 and ∞. The number of clusters can be considered good if the *Dunn* index value is large. The following is how to calculate the *Dunn* index.

$$D(C) = \frac{\min_{C_k, C_l \in C, C_k \neq C_l} \left(\min_{i \in C_k, j \in C_l} d(i,j) \right)}{\max_{C_m \in C} d(C_m)}$$

   with:

   $C = \{C_1, C_{2,} \ldots, C_k\}$ = cluster

   d = Cluster distance measure

   $d(C_m)$ = maximum distance between objects in a cluster C_m

   b. *Silhoueutte* index

   The silhouette index determines the best cluster level by measuring the distance between centroids. The silhouette index has an interval of [-1,1]. The best cluster has a silhouette index value close to 1, so the silhouette index value must be maximized to achieve the best cluster. The method for calculating the Sihoueutte index is as follows:

$$S = \frac{b_i - a_i}{\max\,(a_i, b_i)}$$

   with

   a_i = average distance between i and other observations in the same cluster

   b_i = average distance between i and observations in the nearest neighboring cluster

   c. *Connectivity* index

   Suppose N denotes the total number of observation rows in the dataset, and L denotes the number of observation columns in numerical form. The connectivity index has a value between zero and ∞ and must be minimized. The best cluster has a minimal connectivity index value. The connectivity index is calculated as follows:

$$Conn(C) = \sum_{i=1}^{N} \sum_{j=1}^{L} X_i, nn_{i(j)}$$

   with,

   $nn_{i(j)}$ = nearest neighbor k from observation I

   L =parameters that determine the number of neighbors that influence the *connectivity* measurement
4. Initialize Random Weights

   Select random values for the initial weight vector w_ji (0) with i=1,2,3…,l and j=1,2,3,…,l where l is the number of neurons in the network.
5. Entering input data

   Take sample x from the input space. Where x= x_1,x_2,… x_m with m being the dimension of the input or data space.
6. Calculating distance

The vector distance D_j is calculated from the difference between the initial weight vector w_ij and the input vector x. The neuron with the minimum distance is declared the winning neuron.

7. Weight update
   Updates are made by changing the weight using:
   $$w_{ji}' = w_{ji} + \alpha(x_i - w_{ji})$$
   with $\alpha =$ learning rate
   The learning rate starts with an initial value of α_0 and then decays or decreases with each iteration but does not reach 0. The change in the learning rate value uses the following equation:
   $$\alpha_n = \alpha_0 \left(1 - \frac{n}{iteration_{max}}\right) n = 1,2, \ldots$$
   with
   $n =$ iteration
   $\alpha_0 =$ initial learning rate
8. Continuing
   Check the stopping condition by calculating the difference between the new and old weight vectors. If the difference between the weights is small or even unchanged, the test can be declared converged and the test can be stopped.

Observation of BSF Larvae Growth and Development

The data observed in this study consisted of data on the increase in BSF larva biomass, as well as the individual growth and development of BSF larvae. Individual growth measurements of larvae included the length of BSF larvae, which was measured using graph paper. The development of BSF larvae was observed qualitatively by observing the morphology of the larvae. Larval biomass measurements were performed using the gravimetric method. BSF larval biomass was calculated using the following formula:

$$M = \frac{G}{g} \times m$$

with:
M : BSF larva biomass
G : Total biomass of BSF larvae and media
g : Biomass of BSF larva samples and media
m : Biomass of BSF larva samples

Temperature and Humidity Monitoring

Temperature observations are conducted using a thermometer, air humidity observations are conducted using a hygrometer, and media moisture observations are conducted using a soil moisture sensor.

Data Analysis Techniques

The data obtained was tabulated in tables and graphs to make it easier for readers to understand and draw conclusions. The tabulated data was analyzed using descriptive tests. The data was analyzed using ANOVA tests with the help of SPSS software. If significant results were obtained with a significant level ≤ 5%, further testing was conducted using the DMRT (Duncan's Multiple Range Test).

Results and Discussion

Data normalization

The following is an example of X1 data standardization:

Using Rstudio software, the average value of X1 is 62.44, and the standard deviation is 24.8167. Thus, the value obtained is

$$X_1' \text{ Simelue} = \frac{X_1 - \overline{X_1}}{s} = \frac{70 - 62.44}{24.8167} = 0.304$$
$$X_1' \text{ Aceh Singkil } = \frac{X_1 - \overline{X_1}}{s} = \frac{82 - 62.44}{24.8167} = 0.788$$
$$X_1' \text{ Aceh Selatan} = \frac{X_1 - \overline{X_1}}{s} = \frac{65 - 62.44}{24.8167} = 0.103$$

Determine the number of clusters

In determining the optimal number of clusters, this study used three cluster validation methods included in internal validation, namely the Dunn index, Silhouette index, and Connectivity index. The optimal number of clusters is the number of clusters with the smallest Connectivity index value, the closest Dunn index value, and the largest Silhouette index value. With the help of R Studio software, the cluster validation results are shown in Table 2.

Table 2. Cluster Validation Comparation

Validation Method	Cluster Amount				
	5	6	7	8	9
Index *Dunn*	0.0600	0.0720	0.0357	0.0478	0.0275
Index *Silhouette*	0.2618	0.2658	0.2358	0.2230	0.2187
Index *Connectivity*	195.3611	192.7171	254.8937	243.4560	263.1087

The cluster validation results in Table 2 show that cluster 6 has the smallest Connectivity index value with a total value of 192.7171. The largest Dunn index value is in cluster 6 with a total value of 0.072, and the largest silhouette index value is in cluster 6 with a total value of 0.2658. Thus, cluster validation using internal validation produces cluster 6 as the best number of clusters.

Clustering using the SOM algorithm

Using RStudio software, cluster results were obtained in the form of a fan diagram as shown in Figure 2.

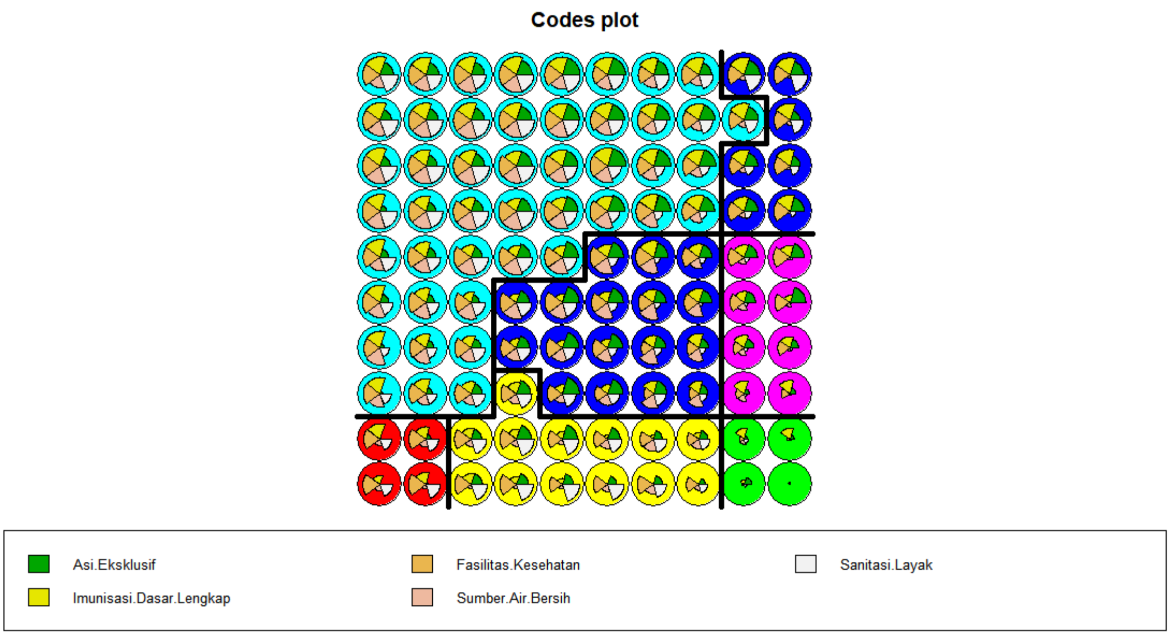

Figure 2. Cluster Results of Fan Diagram

In Figure 2, the fan diagram shows clusters separated by lines with six different colors, each color representing a cluster. Red represents cluster 1, yellow represents cluster 2, green represents cluster 3, light blue represents cluster 4, dark blue represents cluster 5, and purple represents cluster 6. The circles in the diagram represent 5 variables, with each color representing each variable. Green represents the variable of exclusive breastfeeding, yellow represents the variable of complete basic immunization, orange represents the variable of health facilities, pink represents the variable of clean water sources, and white represents the variable of proper sanitation. The characteristics of each cluster are described in Table 3.

Table 1. Cluster Characteristics

Variable	Cluster						Overal Data
	1	2	3	4	5	6	
ASI exclusive (X_1)	11.67	59.18	23.87	68.7	72.29	56.5	62.43
Base vaccination (X_2)	66.65	41.9	43.98	88.09	76.5	65.3	74.65
Health facility (X_3)	94.45	84.17	19.87	95.28	84.9	74.1	86.95
Water dringking source (X_4)	33.94	40.92	10.06	82.68	57.9	27.1	63.01
Proper sanitation (X_5)	65.15	77.98	11.7	78.53	56.6	18.6	65.90

Table 3 shows that each cluster has different characteristics. The lowest average value of variable X_1 is in cluster 1, while the highest average value is in cluster 5. This is consistent with Figure 2, where variable X_1 in cluster 1, represented by the color green, has a smaller portion than the green portion in cluster 5. The lowest average value of variable X_2 in Table 3 is in cluster 2 and the highest is in cluster 4. When viewed in Figure 2, the yellow section representing variable X_2 in cluster 2 is smaller than the yellow section in cluster 4. The lowest average value of variable X_3 in Table 3 is in cluster 3 and the highest is in cluster 4. In Figure 2, variable X_3 is represented by the orange section, and the orange section in cluster 4 is larger than the orange section in cluster 3. The lowest average value of variable X_4 in Table 3 is in cluster 3, and the highest is in cluster 4. In Figure 2, variable X_4, represented by the pink color in cluster 3, has a smaller portion than cluster 4. The lowest average value of variable X_5 in Table 3 is in cluster 3, and the highest average value is in cluster 4. In Figure 2, variable X_5, represented by the white portion, has a larger portion in cluster 4 than in cluster 3.

Table 4. Cluster Result

Kabupaten/Kota	Cluster
Aceh Barat Daya, Kepulauan Mentawai, Pesisir Selatan, Solok, Sijunjung, Tanah Datar, Padang Pariaman, Agam, Lima Puluh, Pasaman, Solok Selatan, Dharmasraya, Pasaman Barat, Padang, Solok, Sawahlunto, Padang Panjang, Bukittinggi, Payakumbuh, Pariaman, Jambi, Tanjab Timur, Ogan Ilir, Mahakam Ulu, Bone Bolango, Gorontalo Utara, Ambon.	1
Aceh Timur, Aceh Barat, Aceh Besar, Pidie, Bereun, Aceh Utara, Aceh Jaya, Pidie Jaya, Bangka, Belitung, Bangka Barat, Bangka Tengah, Bangka Selatan, Belitung Timur, Pangkalpinang, Pandeglang, Sumba Barat, Sumba Timur, Kupang, Timor Tengah Selatan, Timor Tengah Utara, Belu, Alor, Lembata, Flores Timur, Sikka, Ende, Manggarai, Rote Ndao, Manggarai Barat, Sumba Barat Daya, Nagekeo, Manggarai Timur, Sabu Raijua, Malaka, Kota Kupang, Kubu Raya, Kota Pontianak, Kota Singkawang, Seruyan, Nunukan, Buton Utara, Kota Tomohon, Halmahera Barat, Halmahera Utara, Halmahera Selatan, Halmahera Timur, Halmahera Tengah, Kepulauan Sula, Pulau Taliabu, Ternate, Tidore Kepulauan, Puncak,Teluk Wondama, Manokwari, Sorong Selatan, Raja Ampat, Tambrauw, Manokwari Selatan	2
Banggai Laut, Buru, Seram Bagian Barat, Seram Bagian Timur, Buru Selatan, Nabire, Biak Numfor, Paniai, Puncak Jaya, Asmat, Yahukimo, Pegunungan Bintang, Tolikara, Waropen, Mamberamo Raya, Nduga, Lanny Jaya, Mamberamo Tengah, Yalimo, Dogiyai, Intan Jaya, Deiyai, Pegunungan Arfak	3

Aceh Tenggara, Aceh Tengah, Aceh Tamiang, Bener Meriah, Banda Aceh, Sabang, Langsa, Tapanuli Selatan, Toba Samosir, Labuhan Batu, Asahan, Simalungun, Dairi, Karo, Deli Serdang, Langkat, Humbang Hasundutan, Samosir, Serdang Bedagai, Batu Bara, Labuhanbatu Selatan, Labuanbatu Utara, Tanjungbalai, Pematang siantar, Tebing Tinggi, Medan, Binjai, Kuantan Singingi, Indragiri Hulu, Pelalawan, Siak, Kampar, Rokan Hulu, Bengkalis, Rokan Hilir, Kepulauan Meranti, Pekanbaru, Dumai, Natuna, Batam, Karimun, Bintan, Lingga, Tanjung Pinang, Anambas, Sungai Penuh, Kerinci, Bungo, Sarolangun, Tebo, Batang Hari, Merangin, Muaro Jambi, Tanjab Barat, Bengkulu Tengah, Ogan Komering Ulu, Ogan Komering Ilir, Lahat, Musi Rawas, Banyuasin, Ogan Komering Ulu Timur, Pali, Palembang, Lubuk Linggau, Tanggamus, Lampung Selatan, Lampung Timur, Lampung Tengah, Lampung Utara, Way Kanan, Tulang Bawang, Pesawaran, Pringsewu, Mesuji, Tulang Bawang Barat, Tangerang, Kota Tangerang, Cilegon, Kota Serang, Tangerang Selatan, Bogor, Sukabumi, Cianjur, Bandung, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pangandaran, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, Kota Banjar, Kep. Seribu, Jakarta Selatan, Jakarta Timur, Jakarta Barat, Jakarta Utara, Kabupaten Cilacap, Kabupaten Banyumas, Kabupaten Purbalingga, Kabupaten Kebumen, Kabupaten Purworejo, Kabupaten Magelang, Kabupaten Boyolali, Kabupaten Klaten, Kabupaten Sukoharjo, Kabupaten Wonogiri, Kabupaten Karanganyar, Kabupaten Sragen, Kabupaten Grobogan, Kabupaten Blora, Kabupaten Rembang, Kabupaten Pati, Kabupaten Kudus, Kabupaten Jepara, Kabupaten Demak, Kabupaten Semarang, Kabupaten Temanggung, Kabupaten Kendal, Kabupaten Batang, Kabupaten Pekalongan, Kabupaten Pemalang, Kabupaten Tegal, Kabupaten Brebes, Kota Magelang, Kota Surakarta, Kota Salatiga, Kota Semarang, Kota Pekalongan, Kota Tegal, Kulon Progo, Gunung Kidul, Sleman, Bantul, Kota Yogyakarta, Kabupaten Pacitan, Kabupaten Ponorogo, Kabupaten Trenggalek, Kabupaten Tulungagung, Kabupaten Blitar, Kabupaten Kediri, Kabupaten Malang, Kabupaten Lumajang, Kabupaten Jember, Kabupaten Banyuwangi, Kabupaten Pasuruan, Kabupaten Sidoarjo, Kabupaten Mojokerto, Kabupaten Jombang, Kabupaten Nganjuk, Kabupaten Madiun, Kabupaten Magetan, Kabupaten Ngawi, Kabupaten Bojonegoro, Kabupaten Tuban, Kabupaten Lamongan, Kabupaten Gresik, Kabupaten Bangkalan, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu, Jembrana, Tabanan, Badung, Gianyar, Bangli, Karangasem, Buleleng, Denpasar, Lombok Barat, Lombok Tengah, Lombok Timur, Sumbawa, Dompu, Sumbawa Barat, Lombok Utara, Kota Mataram, Kota Bima, Sambas, Tanah Laut, Kotabaru, Banjar, Barito Kuala, Tapin, Hulu Sungai Selatan, Hulu Sungai Tengah, Hulu Sungai Utara, Tabalong, Tanah Bumbu, Balangan, Kota Banjarmasin, Kota Banjarbaru, Kotawaringin Barat, Kotawaringin Timu, Katingan, Barito Timur, Palangkaraya, Kutai Barat, Berau, Penajam Paser Utara, Balikpapan, Samarinda, Bontang, Bulungan, Tana Tidung, Tarakan, Boalemo, Gorontalo, Pohuwato, Kota Gorontalo, Majene, Mamuju, selayar, Bulukumba, Bantaeng, Jeneponto, Takalar, Gowa, Sinjai, Pangkep, Wajo, Sidrap, Pinrang, Luwu, Tana Toraja, Luwu Utara, Luwu Timur, Pare-pare, Palopo, Palu, Bolaang Mongondow, Minahasa, Kepulauan Sangihe, Minahasa Selatan, Minahasa Utara, Bolaang Mongondow Utara, Kepulauan Sitaro, Minahasa Tenggara, Kota Manado, Kota Bitung, Kota Kotamobagu, Merauke, Jayapura,Sorong	4
Simelue, Aceh Singkil, Aceh Selatan, Gayo Luwes, Nagan Raya, Lhoksumawe, Subussalam, Tapanuli Tengah, Tapanuli Utara, Pakpak Bharat, Sibolga, Padang sidimpuan, Gunungsitoli, Indragiri Hilir, Bengkulu Selatan, Rejang Lebong, Bengkulu Utara, Kaur, Seluma, Muko-muko, Lebong, Kepahiang, Bengkulu, Muara Enim, Musi Banyuasin, Empat Lawang, Prabumulih, Lampung Barat , Pesisir Barat, Bandar Lampung, Metro, Lebak, Serang, Garut, Jakarta Pusat, Kabupaten Banjarnegara, Kabupaten Wonosobo, Kabupaten Bondowoso, Kabupaten Situbondo, Kabupaten Probolinggo, Kabupaten Sampang, Kabupaten Pamekasan, Kabupaten Sumenep, Klungkung, Bima, Ngada, Sumba Tengah, Bengkayang, Landak, Mempawah,	5

Sanggau, Ketapang, Sintang, Kapuas Hulu, Sekadau, Melawi, Kayong Utara, Kapuas, Barito Utara, Sukamara,Lamandau, Pulang Pisau, Paser, Kutai Timur, Malinau, Polewali Mandar, Mamasa, Pasangkayu, Mamuju Tengah, Maros, Barru, Bone Bolango, Soppeng, Enrekang, Toraja Utara, Makassar, Buton, Muna, Kolaka, Kolaka Utara, Buton Tengah, Buton Selatan, Kendari, Baubau, Banggai Kepulauan, Banggai, Morowali, Poso, Donggala, Toli-toli, Buol, Parigi Mountog, Tojo Una-una, Sigi, Morowali Utara, Kepulauan Talaud, Bolaang Mongondow Selatan, Bolaang Mongondow Timur, Kepulauan Aru, Fakfak, Kaimana, Teluk Bintuni, Kota Sorong	
Nias, Mandailing Natal, Nias Selatan, Padang Lawas Utara, Padang Lawas, Nias Utara, Nias Barat, Ogan Komering Ulu Selatan, Musi Rawas Utara, Pagar Alam, Barito Selatan, Gunung Mas, Murung Raya, Kutai Kartanegara, Konawe, Konawe Selatan, Bombana, Wakatobi, Konawe Utara, Kolaka Timur, Konawe Kepulauan, Muna Barat, Maluku Tenggara Barat, Maluku Tenggara, Maluku Tengah, Maluku Barat Daya, Tual, Pulau Morotai, Jayawijaya, Kepulauan Yapen, Mimika, Boven Digoel, Mappi, Sarmi, Keerom, Supiori, Kota Jayapura, Maybrat	6

Clustering Results Analysis

Cluster 1 is a group with an average percentage of children under 6 months of age who receive exclusive breastfeeding of 11.67%, an average percentage of children aged 12-23 months who receive complete basic immunization of 66.65%, an average percentage of married women aged 15-49 years who had their last birth with access to health facilities and assistance from health workers of 94.45%, an average percentage of residences with access to proper drinking water sources of 33.94%, and an average percentage of residences with access to proper sanitation of 65.15%.

Cluster 2 is a group with an average percentage of children under 6 months of age who receive exclusive breastfeeding of 59.18%, an average percentage of children aged 12-23 months who receive complete basic immunization of 41.9%, an average percentage of married women aged 15-49 years who had their last birth in a health facility and were assisted by health personnel of 84.17%, an average percentage of dwellings with access to safe drinking water of 40.92%, and an average percentage of dwellings with access to proper sanitation of 77.98%.

Cluster 3 is a group with an average percentage of children under 6 months of age who receive exclusive breastfeeding of 23.87%, an average percentage of children aged 12-23 months who receive complete basic immunization of 43.98%, an average percentage of married women aged 15-49 years who had their last birth in a health facility and were assisted by health personnel of 19.87%, an average percentage of dwellings with access to safe drinking water of 10.06%, and an average percentage of dwellings with access to proper sanitation of 11.7%.

Cluster 4 is a group with an average percentage of children under 6 months of age who receive exclusive breastfeeding of 68.7%, an average percentage of children aged 12-23 months who receive complete basic immunization of 88.09%, an average percentage of married women aged 15-49 years who had their last birth in a health facility and were assisted by health personnel of 95.28%, an average percentage of dwellings with access to safe drinking water of 82.68%, and an average percentage of dwellings with access to proper sanitation of 78.53%.

Cluster 5 is a group with an average percentage of children under 6 months of age who receive exclusive breastfeeding of 72.29%, an average percentage of children aged 12-23 months who receive complete basic immunization of 76.5%, an average percentage of married women aged 15-49 years who had given birth recently and received health facilities and assistance from health workers of 84.9%, an average percentage of residences with access to proper drinking water sources of 57.9%, and an average percentage of residences with access to proper sanitation of 56.6%.

Cluster 6 is a group with an average percentage of children under 6 months of age who receive exclusive breastfeeding of 56.5%, an average percentage of children aged 12-23 months who receive complete basic immunizations of 65.3%, an average percentage of married women aged 15-49 years who gave birth last time with access to health facilities and assisted by health workers of 74.1%, an average percentage of dwellings with access to proper drinking water sources of 27.1%, and an average percentage of dwellings with access to proper sanitation of 18.6%.

Using RStudio software, a map of Indonesia based on clusters is depicted in Figure 3. The red color in Figure 3 represents cluster 1, yellow represents cluster 2, green represents cluster 3, light blue represents cluster 4, dark blue represents cluster 5, and purple represents cluster 6.

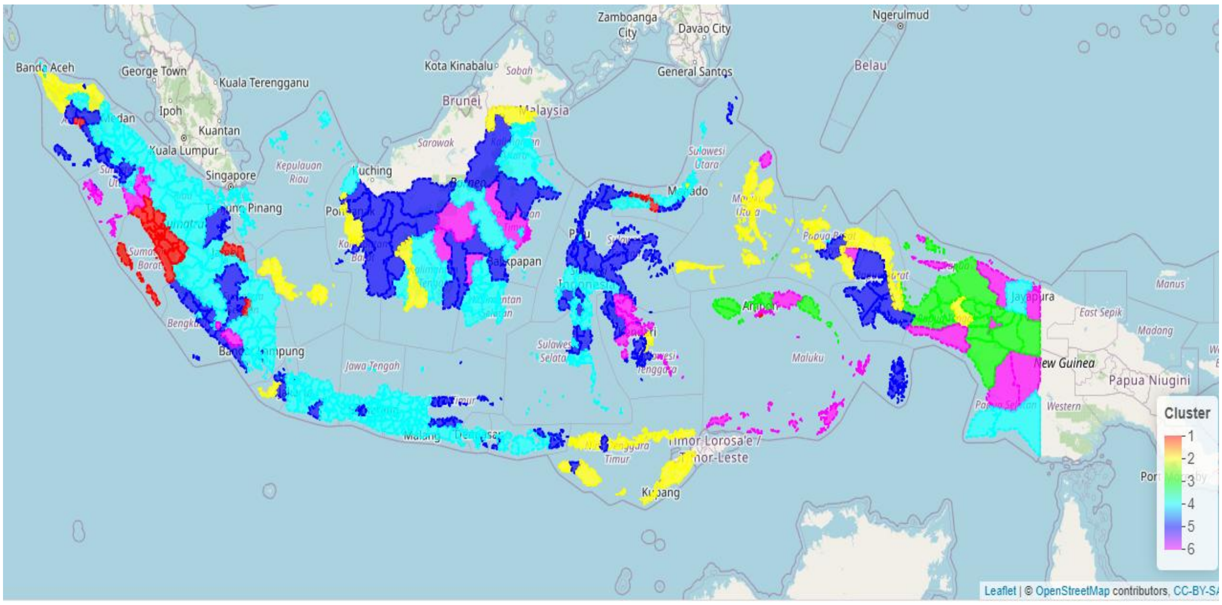

Figure 3 Indonesian Map based on Clusters

Conclusion

The clusters obtained from the analysis using the SOM method numbered 6 clusters. Cluster 1 has 27 districts/cities as members, cluster 2 has 59 districts/cities as members, cluster 3 has 23 districts/cities as members, cluster 4 has 264 districts/cities as members, cluster 5 has 103 districts/cities as members, and cluster 6 has 38 districts/cities as members.

References

Anggraeni, M. R., Yudatama, U., & Maimunah. (2022). Clustering Prevalensi Stunting Balita Menggunakan Agglomerative Hierarchical Clustering. *JURNAL MEDIA INFORMATIKA BUDIDARMA*.

BPS. (2019). *Profil Statistik Kesehatan 2019.* Jakarta: Badan Pusat Statistik.

Firmansyah, D. N., Adinugroho, S., & Rahayudi, B. (2019). Pengelompokan Wilayah Berdasarkan Kesejahteraan Sosial Menggunakan. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (3rd ed.). Waltham: Morgan Kaufmann.

Kasih, P., & Riski, M. (2019). Pemodelan Self Organizing Maps (SOM) Untuk Pengelompokkan Pengaduan Masyarakat pada Radio X. *Generation Journal* .

Kemenkes. (2018). *Buletin Jendela Data dan Informasi Kesehatan.* Jakarta: Pusat Data Kementrian Kesehatan.

Kominfo. (2019). *Bersama Perangi Stunting* . Jakarta: Direktorat Jenderal Informasi dan Komunikasi Publik Kementrian Komunikasi dan Informatika .

Pohan, H., Zarlis, M., Irawan, E., Okprana, H., & Pranayama, Y. P. (2021). Penerapan Algoritma K-Medoids dalam Pengelompokan Balita Stunting di Indonesia. *JUKI : Jurnal Komputer dan Informatika*.

Suwirmayanti, N. L. (2020). Penerapan Teknik Clustering untuk Pengelompokkan Konsentrasi Mahasiswa dengan Metode Self Organizing Map. *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, 11-20.

TNP2K. (2017). *100 Kabupaten/Kota Prioritas untuk Intervensi Anak Kerdil (Stunting).* Jakarta: Tim Nasional Percepatan Penanggulangan Kemiskinan.

Windarto, A. P., Nasution, D., Wanto, A., Tambunan, F., Hasibuan, M. S., Siregar, M. N., . . . Nofriansyah, D. (2020). *Jaringan Saraf Tiruan : Algoritma Prediksi & Implementasi.* (J. Simamarta, Ed.) Yayasan Kita Menulis.