

A Self-Organizing Map Approach for Clustering Provinces Based on Multisectoral Indicators of Stunting Determinants

Admi Salma^{1*}, Riwi Dyah Pangesti², dan Reny Wulandari³

¹Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

²Jurusan Matematika, Universitas Bengkulu, Bengkulu, Indonesia

³Jurusan Manajemen Pertanian lahan Kering, Politeknik Pertanian Negeri Kupang, Kupang, Indonesia

*Corresponding author: admisalma@fmipa.unp.ac.id

Submitted : 14 April 2026

Revised : 10 Mei 2026

Accepted : 30 Mei 2026

ABSTRACT

Stunting is a national issue in Indonesia and also a global challenge. It becomes one of the key priorities outlined in the Sustainable Development Goals (SDGs). The heterogeneity of multisectoral conditions across provinces also contributes to the variation in stunting prevalence in Indonesia. The implementation of uniform policies to address stunting may not yield optimal results due to the diverse needs of each province. Therefore, specific interventions are required to overcome stunting issues. Based on this condition, it is important to cluster provinces based on their characteristics so that the government can determine appropriate interventions for each provincial cluster. Visualization of stunting conditions and multisectoral indicators can also enrich the understanding of each cluster. This study aims to construct clusters of provinces with similar characteristics in terms of multisectoral indicators of stunting determinants. This study applies cluster analysis using a Self-Organizing Map (SOM) algorithm to group provinces. The research steps include data preprocessing, clustering using the SOM algorithm, SOM mapping, and cluster characterization analysis. The results of this study show that three clusters were obtained. The first cluster consists of three provinces characterized by a high maternal mortality rate and a high percentage of exclusive breastfeeding. The second cluster includes nine provinces and is characterized by high risks in maternal and child health as well as economic vulnerability. In addition, the third cluster consists of 26 provinces characterized by relatively good living conditions and quality education.

Keywords: Clustering, Multisectoral, Self-Organizing Map, Stunting.



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Penurunan angka stunting merupakan salah satu fokus utama pemerintah Indonesia saat ini. Komitmen ini tertuang dalam Rencana Pembangunan Jangka Menengah Nasional (RPJMN) yang menargetkan penurunan angka stunting hingga menjadi 14,2% di tahun 2029 (Kementerian Perencanaan Pembangunan Nasional, 2025). Dari hasil Survei Status Gizi Indonesia (SSGI), tercatat bahwa prevalensi stunting Indonesia di tahun 2024 adalah 19,8 %, sehingga pemerintah harus berhasil menurunkan prevalensi stunting sebesar 5,6% dalam 5 tahun kedepan agar target RPJMN tercapai. Hal ini menjadi tantangan nyata mengingat disparitas stunting antar provinsi yang tidak terelakan, beberapa provinsi saat ini memiliki prevalensi stunting tinggi, seperti pada Provinsi Nusa Tenggara Timur (37%), Sulawesi Barat (35,4%), Papua Barat Daya (30,5%). Kondisi ini sangat kontras dengan Bali (8,6%), Jawa Timur (14,7%) dan Kepulauan Riau (15%) (Kementerian Kesehatan, 2025).

Tingginya keragaman prevalensi stunting mengisyaratkan bahwa stunting bukan hanya masalah tunggal, namun berkaitan dengan kondisi multisektoral, seperti kondisi ekonomi, pendidikan, kesehatan dan sosial. Hal ini sejalan dengan beberapa riset terdahulu, yang menekankan bahwa tingginya disparitas stunting dipengaruhi oleh faktor multisektoral seperti faktor ekonomi, sosial, pendidikan orang tua dan erat kaitannya dengan kemiskinan (Allo dkk, 2025 & Azzam dkk, 2024). Kondisi ekonomi dan tekanan finansial juga mempengaruhi kualitas kesehatan anak (Spencer dan Komro, 2017 & Francis dkk, 2018).

Untuk menghadapi kondisi ini, pendekatan yang bersifat umum dan seragam cenderung tidak efektif. Diperlukan langkah yang tepat, sehingga pemerintah dapat memberikan penanganan stunting yang sesuai dengan kebutuhan masing-masing provinsi guna mencapai tujuan penanganan stunting di Indonesia. Pengelompokan provinsi berdasarkan kemiripan karakteristik menjadi salah satu cara efektif.

Pendekatan yang dapat digunakan adalah analisis *cluster*. Analisis *cluster* mampu mengelompokkan objek berdasarkan kemiripan karakteristiknya (Az-Zahra, 2022). Dari penelitian terdahulu, telah banyak yang membahas pengelompokan objek berdasarkan indikator stunting atau kesehatan ibu dan anak. Variabel yang umumnya digunakan adalah Jumlah bayi mendapatkan ASI, Jumlah bayi dengan imunisasi lengkap, dan Jumlah bayi penerima IMD (Meilani dkk, 2023; Taufik dkk, 2023; & Fadilah dkk 2022). Namun, *cluster* yang terbentuk tidak mengakomodir kondisi sosial ekonomi (multi sektoral) yang juga menjadi penyebab terjadinya stunting di Indonesia.

Analisis *cluster* yang sering digunakan adalah pendekatan *Cluster* Hirarki atau K-means *Cluster* baik untuk kasus stunting maupun sosila ekonomi (Anitha dan Patil, 2021; Meilani dkk, 2023; Salma dan Zilrahmi, 2025). Namun, analisis *cluster* tersebut belum mampu memberikan visualisasi data yang menggambarkan kontribusi variabel dalam *cluster* yang terbentuk secara mendalam, sehingga informasi belum dapat dimanfaatkan secara maksimal.

Berdasarkan gap yang ditemukan dari penelitian sebelumnya, maka dalam penelitian ini akan dilakukan pengelompokan provinsi di Indonesia, dengan menambahkan variable multisektoral yang berkontribusi dalam kondisi stunting. Analisis *cluster* melalui pendekatan *Self Organizing Maps* (SOM) dipilih karena selain mampu meng*cluster*kan objek, juga memvisualisasikan data kedalam bentuk dua dimensi sehingga lebih mudah dipahami. Dalam penelitian ini, penerapan SOM bertujuan untuk menghasilkan visualisasi dan pengelompokan provinsi rentang stunting melalui kondisi multisektoral, yang mencakup area kesehatan, pendidikan dan, sosial ekonomi. Sehingga hasil dari penelitian ini dapat dimanfaatkan untuk kepentingan kebijakan dalam membuat intervensi guna mempercepat penyelesaian masalah stunting.

II. METODE PENELITIAN

Penelitian ini adalah penelitian kuantitatif dengan menerapkan metode *cluster* berbasis *Self Organizing Maps* (SOM). SOM merupakan bagian dari *unsupervised learning* dalam data mining, dan mampu memvisualisasikan data dalam bentuk dua dimensi. Beberapa penelitian terdahulu menemukan bahwa SOM juga memberikan performa lebih baik dari metode *cluster* lainnya seperti K-means *cluster* atau fuzzy C-means (Az-Zahra, 2022). Dalam studi ini, SOM digunakan untuk mengelompokkan provinsi di Indonesia berdasarkan faktor penyebab stunting. Data yang digunakan dalam penelitian ini adalah data sekunder yang dirilis oleh kementerian Kesehatan (Kemenkes) dengan judul “Profil Kesehatan Indonesia 2024”, yang dirilis pada tahun 2025. Data yang digunakan melibatkan data dari 38 Provinsi di Indonesia dengan 9 variabel sebagai faktor penyebab terjadinya stunting di Indonesia. Adapun variabelnya adalah Persentase Bayi Berat Badan Lahir Rendah (BBLR) (X1), jumlah kematian ibu (X2), Persentase Bayi Baru Lahir Mendapat Inisiasi Menyusu Dini (IMD) (X3), Bayi mendapat Asi eksklusif 6 bulan (X4), Persentase Rumah Tangga Yang Memiliki Akses Terhadap Sanitasi Layak (X5), rata-rata lama sekolah bagi Perempuan (X6), persentase Pasangan Usia Subur (PUS) terlalu (X7), Persentase penduduk miskin (X8) dan Indeks Pembangunan Manusia (IPM) (X9). PUS terlalu mencakup pasangan usia subur pada rentan usia kurang dari 15 tahun, lebih dari 35 tahun, yang memiliki banyak anak dan jarak kelahiran terlalu dekat.

Adapun penelitian ini akan menggunakan *software* R dengan langkah kerja (Qian dkk, 2019) seperti berikut:

1. Persiapan data

Pada tahap ini dilakukan pengumpulan data. Setelah data terkumpul dilakukan preprocessing data, dimana dilakukan pengecekan data hilang, dan data ganda, serta standarisasi data. Adapun standarisasi data (Z) diperoleh dari perbandingan selisih data (x) dan rata-ratanya (\bar{x}) dengan simpangan baku (σ)

$$Z = \frac{x - \bar{x}}{\sigma}$$

2. *Clustering* dengan algoritma SOM (Yahya dkk, 2024; Putri & Maulidia, 2024)

- a. Menentukan Jumlah *cluster* yang akan dibentuk

- b. Melakukan inisiasi awal SOM

Tahapan ini dimulai dengan inisiasi bobot awal W_{ij} yang dipilih secara acak. Selanjutnya menentukan learning rate yang bernilai 0,1, serta penetapan nilai radius R dalam jaringan kohonen, dimana R bernilai nol (0).

- c. Setelah menentukan bobot awal, learning rate dan nilai R, maka akan dihitung jarak Euclidean yang menggambarkan jarak dari antar bobot ke input vektor xi untuk setiap *cluster* J

$$D(W_{ij}, x_n) = \sqrt{\sum (w_{ij} - x_{ij})^2}$$

- d. Menentukan neuron J pemenang berdasarkan jarak terkecil yang disebut *Best Matcing Unit* (BMU)
- e. Menghitung *neighborhood function* (fungsi tetangga)

$$h_{ci}(t) = \exp\left(\frac{\|r_c - r_i\|^2}{2R_{(t)}^2}\right)$$

Dimana r_c dan r_i merupakan posisi BMU dan neuron ke-i

- f. Memperbaharui bobot neuron pemenang dan neuron tetangga, dimana nilai bobot neuron pemenangnya adalah:

$$W_{ij(t+1)} = W_{ij(t)} + \alpha(x_{ij} - W_{ij(t)})$$

- g. Lakukan iterasi langkah c-e untuk semua data input
3. Melakukan profilisasi hasil *cluster* dan interpretasi
Pada tahap ini, akan ditentukan karakteristik *cluster* dan memaknai masing masing karakteristik *cluster* sebagai penciri masing masing *cluster*.
 4. Melakukan validasi *cluster* setelah hasil *cluster* diperoleh
Dalam penelitian ini, validasi yang digunakan adalah *connectivity*, *Dunn Index*, dan *Silhouette*. Ketiga validasi internal ini digunakan untuk melihat kualitas *cluster* yang terbentuk dan kejelasan posisi data dalam *cluster* sehingga tidak terjadi tumpang tindih data. Semakin rendah nilai *connectivity* mengindikasikan semakin baik *cluster* yang terbentuk. Sementara semakin tinggi nilai *dunn index* dan *silhouette* maka semakin tepat pemisahan dan peletakan anggota *cluster*.

III. HASIL DAN PEMBAHASAN

Penelitian ini melibatkan 9 variabel yang merupakan penyebab stunting di Indonesia. Pada tahap awal, dilakukan eksplorasi data melalui statistik deskriptif.

Tabel 1. Statistik Deskriptif

Variabel	Minimum	Median	Mean	Max
Persentase Bayi Berat Badan Lahir Rendah (BBLR)	0,93	4,11	4,53	8,01
Jumlah Kematian Ibu	4,00	69,00	108,29	749,0
Persentase Bayi Baru Lahir Mendapat Inisiasi Menyusu Dini (IMD)	69,71	92,94	101,96	196,6
Persentase Bayi Mendapat Asi Eksklusif 6 Bulan	10,41	59,70	56,62	83,42
Persentase Rumah Tangga Yang Memiliki Akses Terhadap Sanitasi Layak	12,61	83,36	81,14	96,83
Rata-rata Lama Sekolah (RLS) Perempuan	3,40	8,72	8,60	11,19
Persentase PUS TERLALU	29,96	41,09	41,55	51,92
Persentase Penduduk Miskin	3,80	9,57	10,66	29,66
Indeks Pembangunan Manusia (IPM)	54,43	74,08	73,50	84,15

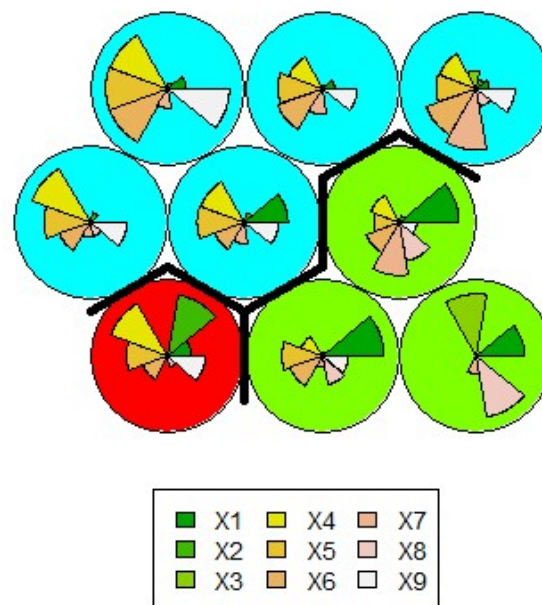
Tabel 1 menyajikan kondisi indikator penyebab stunting di Indonesia. Ketimpangan terjadi pada semua indikator dan menjadi indikasi tingginya disparitas antar provinsi di Indonesia. Beberapa provinsi masih memiliki angka yang tinggi terkait jumlah kematian ibu, PUS terlalu, BBLR dan persentase penduduk miskin. Semakin tinggi angka tersebut, maka kondisi provinsi akan semakin jauh dari keadaan ideal yang ditargetkan pemerintah. Sementara, rendahnya angka rata-rata lama sekolah bagi, akses sanitasi layak, dan Bayi mendapat Asi Eksklusif 6 bulan di beberapa provinsi juga

menandakan bahwa kondisi tersebut masih harus menjadi perhatian pemerintah agar dapat diintervensi. Dari Tabel 1 juga diketahui bahwa persentase tertinggi IMD di Indonesia mencapai angka 196,6 %. Hal ini terjadi karena perbedaan sumber data untuk numerator dan denominator, sehingga diperoleh informasi jumlah IMD lebih besar dari jumlah kelahiran bayi, perbedaan data ini menyebabkan persentase IMD lebih dari 100% (Rahayuningsih dan Sufriani, 2025).

Langkah selanjutnya adalah melakukan *cluster* pada data yang telah distandarisasi. Pada penelitian ini digunakan grid 3x3, untuk menghasilkan 3 *cluster*. Adapun hasil *cluster* melalui pendekatan SOM disajikan pada Tabel 2, dan divisualisasikan pada Gambar 1.

Tabel 2. Anggota Cluster

Cluster	Jumlah anggota	Provinsi
1	3	Jawa Barat, Jawa Tengah, Jawa Timur
2	9	Nusa Tenggara Timur, Gorontalo, Sulawesi Barat, Papua, Papua Barat, Papua Selatan, Papua Tengah, Papua Pegunungan, Papua Barat Daya
3	26	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, DIY Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Tengah, Sulawesi Utara, Sulawesi Selatan, Sulawesi Tenggara, Maluku, Maluku Utara



Gambar 1. Unified Distance Matrix

Gambar 1 menunjukkan 3 *cluster* yang terbentuk melalui pendekatan SOM, dimana masing-masing *cluster* diwakili oleh warna merah (1 neuron/lingkaran), Hijau (3 neuron) dan toska (5 neuron). Jumlah neuron tidak merepresentasikan jumlah objek pada tiap *cluster*. *Cluster* merah (*cluster* 1) memiliki 3 anggota, yaitu Jawa Barat, Jawa Tengah, dan Jawa Timur. Dibandingkan dengan *cluster* lain, *cluster* ini memiliki angka kematian ibu (X2) dan pemberian asi eksklusif (X4) lebih tinggi, sementara variabel lain bukanlah penciri dari *cluster* ini. *Cluster* ini mengindikasikan tingginya kematian ibu, namun praktik ASI baik dibandingkan *cluster* lain. Pada *cluster* ini, 74,36% bayi telah mendapatkan ASI eksklusif, sementara jumlah kematian ibu yang terjadi setara dengan 89 kematian ibu dari 100.000 kelahiran hidup. Walaupun jumlah kematian ibu adalah yang tertinggi di *cluster* ini, namun menurut RPJMN angka ini sudah berada

dibawah target Nasional untuk Angka kematian Ibu (Kementerian Perencanaan Pembangunan Nasional, 2025), yaitu sebesar 189 kematian ibu dari 100.000 kelahiran hidup.

Cluster selanjutnya adalah *cluster* kedua (neuron hijau) yang terdiri dari 9 provinsi. Pada *cluster* ini memiliki karakteristik persentase BBLR (X1), bayi mendapatkan IMD (X3), PUS terlalu (X7) dan Persentase penduduk Miskin (X8) yang lebih tinggi dari *cluster* lainnya. Persentase BBLR dan PUS terlalu yang tinggi menjadi indikasi bahwa dalam *cluster* ini memiliki indikator kondisi kesehatan rentan untuk bayi dan ibu yang tinggi. Sementara tingginya persentase kemiskinan juga menjadi indikasi kerentanan kondisi sosial ekonomi. Namun *cluster* ini merupakan *cluster* dengan bayi yang mendapatkan IMD tertinggi. Sehingga *cluster* ini memiliki indikasi kerentanan ekonomi dan kesehatan ibu dan bayi, dengan praktik IMD yang baik.

Selanjutnya, *cluster* ketiga memiliki 26 anggota. *Cluster* ini memiliki rumah dengan sanitasi layak (X5), Persentase RLS perempuan (X6) dan IPM (X9) yang lebih tinggi dibanding dengan dua *cluster* lainnya. Hal ini mengindikasikan bahwa penghidupan yang layak dan pendidikan untuk perempuan sudah baik di *cluster* tersebut.

Selanjutnya dilakukan uji validasi *cluster* guna mengukur ketepatan pembagian *cluster* yang telah dilakukan.

Tabel 3. Validasi *Cluster*

Validasi <i>Cluster</i>	Jumlah <i>Cluster</i>				
	3	4	5	6	7
<i>Connectivity</i>	9,58	17	23,23	34,65	41,38
<i>Dunn</i>	0,156	0,069	0,0645	0,049	0,049
<i>Silhouette</i>	0,462	0,422	0,396	0,296	0,28

Tabel 3 menyajikan nilai validasi *cluster* untuk 3 *cluster* yang terbentuk hingga 7 *cluster* yang terbentuk. Dari informasi tersebut, terlihat bahwa jumlah *cluster* terbaik yang dapat dibentuk dalam penelitian ini adalah 3 *cluster* dari semua opsi kluster yang tersedia. Hal ini karena nilai *connectivity* terendah, nilai *Dunn* tertinggi dan nilai *Silhouette* tertinggi. Sehingga, metode SOM yang diterapkan dapat melakukan pengelompokan data optimal pada 3 *cluster*

IV. KESIMPULAN

Dalam penelitian ini dilakukan pengelompokan provinsi melalui pendekatan SOM berdasarkan indikator penyebab stunting baik dari segi kesehatan hingga sosial ekonomi. Dari pengelompokan yang telah dilakukan, terbentuk 3 *cluster* dengan karakteristik yang berbeda. *Cluster* pertama memiliki karakteristik tingginya angka kematian Ibu, namun Praktik ASI eksklusif 6 bulan sudah terlaksana dengan baik. Sementara *cluster* kedua terdiri dari 9 provinsi dengan karakteristik kerentanan ekonomi dan kesehatan ibu dan bayi, namun praktik IMD baik. *Cluster* 3 terdiri dari 26 provinsi dengan penciri kehidupan layak dan pendidikan bagi perempuan yang lebih baik dibandingkan dua *cluster* lainnya. Sehingga, dari 3 *cluster* yang terbentuk, maka *cluster* yang sangat rentan adalah *cluster* kedua. Selanjutnya, dari hasil validasi *cluster* diperoleh SOM mampu membentuk 3 *cluster* optimal dari data yang digunakan.

Berdasarkan penelitian yang telah dilakukan, *cluster* kedua menjadi *cluster* paling rentan dan butuh perhatian dan intervensi dari pemangku kebijakan. Dalam *cluster* tersebut terjadi kerentanan kesehatan ibu dan kemiskinan yang tinggi, sehingga disarankan untuk penelitian selanjutnya untuk mengkaji terkait hal tersebut, baik dengan melakukan penambahan variabel seperti pelayanan kesehatan yang diterima ibu, jumlah tenaga medis dan lainnya.

DAFTAR PUSTAKA

- Allo, C.B.G., Nicea, R. P., Winda, A. F. B., Bobi, & F. K., Feby, S. (2025), Analysis of Stunting Data in Indonesia Using K-Means and Self Organizing Map (SOM). *Statistika*, 25(2). <https://journals.unisba.ac.id/index.php/statistika/article/view/7778>
- Anitha, P., & Patil, M.M. (2021). RFM model for customer purchase behavior using k-means algorithm. *Journal of King Saud University-Computer and Information Sciences*, 34,1785-1792
- Az-Zahra, R. (2022). Analisis Perbandingan Metode Self Organizing Map dengan Metode Fuzzy C-Mean pada Pengelompokan Peminatan Jurusan di Sekolah Kejuruan. *Multitek Indonesia*, 16(2), 65-74
- Azzam, R. M., Firmansyah, G., Mulyo, W. A., & Tjahjono B. (2024). Clustering of Child Stunting Data in Tangerang Regency Using Comparison of K-Means, Hierarchical Clustering and DBSCAN Methods. *Asian J Soc Humanit*, 2(12).

- Francis L, DePriest K, Wilson M, Gross D. (2018). Child Poverty, Toxic Stress, and Social Determinants of Health: Screening and Care Coordination. *OJIN Online J Issues Nurs*, 23(3).
- Fadilah, A., Pangestu, M.N., Lumbanbatu, S., & Defiyanti, S. (2022). PENGELOMPOKAN KABUPATEN/KOTA DI INDONESIA BERDASARKAN FAKTOR PENYEBAB STUNTING PADA BALITA MENGGUNAKAN ALGORITMA K-MEANS. *JIKO J Inform Dan Komputer*, 6(2):223.
- Meilani, D., Masjkur, M., & Afendi, F.M.(2023). Grouping Provinces in Indonesia Based on the Causes of Stunting Variables using Hierarchical *Clustering* Analysis: Pengelompokan Provinsi di Indonesia Berdasarkan Peubah Penyebab Stunting Menggunakan Analisis *Cluster* Hierarki. *Indones J Stat Its Appl (IJSA)*,7(1):32–43.
- Taufik, H., Mohamad, J., & Susilawati. (2023). *Clustering* daerah rawan stunting di Jawa Barat menggunakan algoritma K-Means. *INFOTECH J Inform Teknol*, 4(2):137–46.
- Kemertian Perencanaan Pembangunan Nasional. (2025) Ringkasan Rencana Pembangunan Jangka Menengah Nasional Tahun 2025-2029.
https://perpustakaan.bappenas.go.id/elibrary/file_upload/koleksi/dokumenbappenas/konten/Dokumen%202025/Konten/%7BDigital%7D%20Ringkasan%20RPJMN%20Tahun%202025-2029.pdf
- Kementerian Kesehatan. (2025). Profil Kesehatan Indonesia 2024. Kementerian Kesehatan Republik Indonesia
- Putri, A. D., Maulidia, S. (2024). Pengelompokan Kejadian Stunting di Indonesia pada Tahun 2022 dan Faktor-faktor yang Mempengaruhinya: Sebuah Gambaran. *Semin Nas Off Stat*,(1):449–5.
- Qian, J., Nguyen, N.P., Oya, Y., Kikugawa, G., Okabe, T., Huang, Y., dkk. (2019). Introducing self-organized maps (SOM) as a visualization tool for materials research and education. *Results Mater*, 4
- Rahayuningsih, S.I., Sufriani, S. (2025). Analisis Korelasi Kunjungan Neonatal, Inisiasi Menyusu Dini, dan Akses Fasilitas Kesehatan terhadap Cakupan ASI Eksklusif di 38 Provinsi Indonesia Tahun 2024. *J Penelit Inov*, 5(4):3383–92
- Salma, A., Zilrahmi, Z. (2025), Multidimensional Poverty *Clustering* using K-Means Algorithm with Dimensionality Reduction by Principal Component Analysis. *Rangkiang Math J*, 4(2):111–8
- Spencer, R.A., Komro, K.A. (2017). Family Economic Security Policies and Child and Family Health. *Clin Child Fam Psychol Rev*, 20(1):45–63
- Yahya, L. M., Kertanah, K., & Hidayaturrehman, U. (2024). Penerapan Algoritma Self Organizing Maps (SOM) dan K-means Untuk Mengelompokan Akseptor KB di NTB. *J Stat Dan Komputasi*, 3(1): 32-41.