

Grouping Level of Poverty Based on District/City in Indonesia Using K-Harmonic Means

Nabillah Putri Shalsabila, Nonong Amalita*, Dodi Vionanda, Dony Permana

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

*Corresponding author: nongaditya@gmail.com

Submitted : 05 April 2023

Revised : 05 Mei 2023

Accepted : 19 Mei 2023

ABSTRACT

Indonesia still has a relatively high poverty rate, although nationally it has declined in recent years. There are areas that are still experiencing increasing poverty rates. So that the currently planned poverty alleviation plans are no longer uniform, but need to pay attention to the conditions of each dimension that cause poverty in an area, so it is necessary to group districts/cities in Indonesia on poverty. Grouping was performed using K-Harmonic Means analysis. K-Harmonic Means is a non-hierarchical clustering that takes the average of the harmonic distance between each data point and the cluster's center. The data used in this research is secondary data sourced from BPS publications on poverty and inequality in 2022. The analysis technique is carried out by standardizing the data, conducting cluster analysis, and validating clusters. Based on the results of the K-Harmonic Means analysis, the optimal number of clusters is two clusters that first cluster has 54 districts/cities while second cluster has 460 districts/cities and the Dunn Index value for cluster validation is 0,03492. So that a better grouping level of poverty based on district/city in Indonesia is obtained by using the K-Harmonic Means method with $p = 2,25$.

Keywords: *Dunn Index, K-Harmonic Means, Poverty Level, Silhouette Coefficient*



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

Tingkat kemiskinan merupakan ukuran yang digunakan untuk mengukur proporsi masyarakat yang berada di bawah garis kemiskinan. Garis kemiskinan ialah nilai rupiah pengeluaran yang diperlukan seseorang untuk memenuhi kebutuhan pokok hidupnya selama sebulan, baik kebutuhan makanan maupun non makanan (BPS, 2022). Tingkat kemiskinan ini dapat diukur dengan menggunakan beberapa indikator seperti persentase penduduk miskin, angka kemiskinan, atau indeks kemiskinan. Badan Pusat Statistik di Indonesia menggunakan persentase penduduk miskin sebagai indikator untuk mengukur tingkat kemiskinan di Indonesia (BPS, 2022).

Persentase penduduk miskin di Indonesia mengalami penurunan mencapai 9,54 persen pada bulan Maret tahun 2022 (BPS, 2022). Menurut Adhitya (2022) Penurunan persentase penduduk miskin tersebut belum mewakili persentase penduduk miskin yang ada di setiap daerah di Indonesia. Pada kenyataannya masih banyak daerah-daerah yang mengalami kenaikan jumlah penduduk miskin yang cukup tinggi, seperti pada Provinsi Sulawesi Barat, Kepulauan Riau, Kalimantan Tengah, Bali, dan Papua (BPS, 2022). Hal ini disebabkan oleh banyak dari kebijakan pemerintah dalam menanggulangi kemiskinan pada setiap daerah di Indonesia disamaratakan saja sehingga pada beberapa daerah kebijakan pengetasan kemiskinan tersebut berjalan tidak tepat sasaran (Arifin, 2020).

Oleh karena itu, agar kebijakan-kebijakan pemerintah dalam menanggulangi kemiskinan di setiap daerah tepat sasaran dan tersebar merata, pemerintah memerlukan pengelompokkan tingkat kemiskinan di setiap daerahnya, sehingga pemerintah mengetahui daerah mana saja yang memiliki tingkat kemiskinan tinggi atau rendah (Putra dan Lisna, 2020). Oleh karena itu, pengelompokkan tingkat kemiskinan setiap daerah di Indonesia dapat dikelompokkan dengan menggunakan analisis *cluster*. Menurut Hermawati (2013:125-127) analisis *cluster* adalah upaya untuk mengelompokkan observasi yang memiliki kesamaan karakteristik tertentu sehingga menjadi kelompok-kelompok.

Dalam penelitian Gunadi dkk., (2019) yakni membandingkan kinerja metode *K-Means* dan *K-Harmonic Means* pada pengelompokkan siswa Bintara di Bali di dapat hasil pada metode *K-Harmonic Means* memiliki jumlah ketidaktepatan pada pengelompokkan lebih kecil dibanding dengan metode *K-Means*. Dalam penelitian (Christienova dkk., 2017) melakukan pengelompokkan desa-desa di Kabupaten Lanny Jaya berdasarkan Data Potensi Desa (DPD)

dengan menggunakan pendekatan *K-Means* dan *K-Harmonic Means* dan hasilnya adalah *K-Harmonic Means* lebih baik dibanding dengan *K-Means*.

Dalam penelitian ini, pendekatan *K-Harmonic Means* digunakan untuk mengelompokkan tingkat kemiskinan kabupaten/kota yang ada di Indonesia, sehingga dapat menjadi salah satu acuan untuk menerapkan kebijakan dalam menanggulangi kemiskinan pada setiap daerah di Indonesia dan juga dapat menjadi rujukan untuk penelitian selanjutnya.

II. METODE PENELITIAN

Penelitian ini merupakan penelitian terapan. Dimana penelitian terapan ini bertujuan untuk memecahkan suatu masalah sehingga hasil dari penelitian ini dapat dimanfaatkan oleh individu atau kelompok untuk kepentingan industri, pemerintah, atau politik dan bukan hanya untuk menambah wawasan saja (Trianto, 2010: 168). Data yang digunakan pada penelitian ini adalah data sekunder yang berasal dari *website* publikasi Badan Pusat Statistik (BPS) tentang kemiskinan dan ketimpangan di Indonesia tahun 2022.

Adapun variabel yang digunakan pada penelitian ini adalah Persentase Penduduk Miskin, Indeks Kedalaman Kemiskinan, Indeks Keparahan Kemiskinan, dan Garis Kemiskinan. Variabel pada penelitian ini ditampilkan pada Tabel 1.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Skala	Satuan
X ₁	Persentase Penduduk Miskin	Rasio	Persem
X ₂	Indeks kedalaman kemiskinan	Rasio	Persen
X ₃	Indeks keparahan kemiskinan	Rasio	Persen
X ₄	Garis Kemiskinan	Rasio	Rupiah/Kapita/Bulan

(Sumber : BPS)

Pada penelitian ini dilakukan analisis data dengan menggunakan algoritma *K-Harmonic Means*. Algoritma *K-Harmonic Means* merupakan teknik pengelompokan yang berbasis pusat membentuk *cluster* dengan memperbaiki setiap *cluster* secara iteratif berdasarkan lokasi titik. *K-Harmonic Means* juga disebut sebagai algoritma pengelompokan yang berupaya meminimalisir rata-rata harmonik dari semua titik data ke pusat *cluster* (Hamerly dan Elkan, 2002). Pada algoritma *K-Harmonic Means* jumlah *cluster* yang digunakan ditentukan dengan melihat nilai *silhouette coefficient*. *Silhouette coefficient* merupakan gabungan dari metode kohesi dan separasi. Nilai kohesi merupakan rata-rata jarak data ke-*i* terhadap *centroid* pada *cluster* itu sendiri. Sedangkan nilai separasi adalah minimum rata-rata jarak data ke-*i* terhadap *centroid* lain pada *cluster* yang berbeda. Nilai *silhouette coefficient* ini memiliki rentang dari -1 hingga 1. Nilai 1 pada *silhouette coefficient* berarti seluruh objek berada pada *cluster* yang tepat, sedangkan nilai -1 berarti bahwa objek tersebut termasuk pada *cluster* yang tidak tepat. Nilai *silhouette coefficient* tertinggi menunjukkan jumlah kelompok yang optimum digunakan dan pengelompokan yang paling baik. Nilai *silhouette coefficient* dihitung dengan menggunakan rumus berikut.

$$SC(i) = \frac{1}{n_k} \sum_{x_{ij} \in C_k} \frac{b_i - a_i}{\max\{a_i, b_i\}}$$

Dengan $i = 1, 2, \dots, n_k$; $k = 1, 2, \dots$; x_{ij} = objek ke-*i* variabel ke-*j*; C_k adalah *cluster* ke-*k*; a_i = nilai kohesi objek ke-*i*; b_i = nilai separasi objek ke-*i*. nilai kohesi dan nilai separasi dapat dihitung dengan menggunakan rumus sebagai berikut (Wierzchon dan Klopotek, 2018:171)..

$$a_i^j = \frac{1}{m_j - 1} \sum_{\substack{r=1 \\ r \neq i}}^{m_j} d(x_i^j, x_r^j)$$

Dengan

j : *cluster*

i : index data

a_i^j : rata-rata jarak data ke-*i* terhadap semua data dalam satu *cluster*

M_j : jumlah data dalam *cluster* ke-*j*

$d(x_i^j, x_r^j)$: jarak data ke-*i* dengan data ke-*r* dalam satu *cluster* *j*

$$b_i^j = \min_{\substack{n=1, \dots, k \\ n \neq j}} \left\{ \frac{1}{m_n} \sum_{\substack{r=1 \\ r \neq i}}^{m_n} d(x_i^j, x_r^j) \right\}$$

Dengan

- b_i^j : rata-rata jarak data ke- i terhadap semua data yang tidak dalam satu *cluster* dengan data ke- i
 M_n : jumlah data dalam *cluster* ke- n
 $d(x_i^j, x_r^j)$: jarak data ke- i dengan data ke- r dalam satu *cluster* j

Menurut Hamerly dan Elkan (2022) tahapan yang dilakukan selanjutnya adalah penentuan titik awal *centroid* ditentukan secara *random*. Setelah itu dilakukan perhitungan nilai fungsi tujuan pada algoritma *K-Harmonic Means* dengan menggunakan parameter ($p \geq 2$) berdasarkan rumus sebagai berikut.

$$KHM(X, C) = \sum_{i=1}^N \frac{K}{\sum_{j=1}^K \frac{1}{\|x_i - c_j\|^p}} \quad (1)$$

Kemudian dilakukan perhitungan nilai keanggotaan pada setiap data dari setiap pusat *cluster* dengan rumus sebagai berikut.

$$m(c_j|x_i) = \frac{\|x_i - c_j\|^{-p-2}}{\sum_{j=1}^K \|x_i - c_j\|^{-p-2}} \quad (2)$$

Setelah itu dilakukan perhitungan nilai bobot setiap data dengan menggunakan rumus sebagai berikut.

$$w(x_i) = \frac{\sum_{j=1}^K \|x_i - c_j\|^{-p-2}}{(\sum_{j=1}^K \|x_i - c_j\|^{-p})^2} \quad (3)$$

Setelah didapat nilai keanggotaan dan nilai bobot pada setiap data maka selanjutnya dapat dilakukan perhitungan kembali *centroid* baru pada setiap titik pusat *cluster* dari seluruh data menggunakan nilai keanggotaan serta bobot pada tiap data dengan rumus sebagai berikut.

$$c_j = \frac{\sum_{i=1}^n m(c_j|x_i) \cdot w(x_i) \cdot x_i}{\sum_{i=1}^n m(c_j|x_i) \cdot w(x_i)} \quad (4)$$

Dengan

- c_j : *cluster* ke- j
 x_i : data ke- i
 $m(c_j|x_i)$: nilai keanggotaan data ke- i pada *cluster* ke- j
 $w(x_i)$: nilai bobot pada data ke- i

Perhitungan pada langkah kedua hingga kelima secara berulang hingga nilai fungsi tujuan tidak berubah secara signifikan. Setelah iterasi berhenti data dialokasikan ke *cluster* menggunakan nilai keanggotaan yang terbesar. Setelah didapatkan hasil pengelompokkan maka dilakukan validasi hasil *cluster* dengan menggunakan *Dunn Index*. *Dunn Index* merupakan metode validasi *cluster* yang melakukan validasi *cluster* berdasarkan informasi internal yang terdapat pada data. *Dunn Index* merupakan validasi *cluster* dihitung dengan cara mengukur jarak antara dua *cluster* dengan jarak maksimum dalam *cluster*. Semakin besar nilai *Dunn Index* maka *cluster* yang terbentuk semakin baik (Wierzchon dan Kłopotek, 2018:171). Perhitungan *Dunn Index* dapat dirumuskan sebagai berikut.

$$D(k) = \min_{i=1, \dots, k} \left\{ \min_{j=i+1, \dots, k} \left(\frac{d(c_i, c_j)}{\max_{v=1} diam(c_v)} \right) \right\} \quad (5)$$

Dengan

$D(k)$: Nilai *Dunn Index*

k : Jumlah *cluster*

$d(c_i, c_j)$: Jarak antara *cluster* i dan *cluster* j

$diam(c_v)$: Diameter *cluster* maksimum

Adapun teknik analisis data yang digunakan pada penelitian ini adalah dimulai dengan menginput data dan menyajikan deskripsi data dari masing-masing variabel yang digunakan. Selanjutnya menentukan jumlah *cluster* dengan menggunakan nilai *silhouette coefficient* dan menentukan titik awal *centroid* secara acak dari masing-masing *cluster*. Selanjutnya menghitung nilai fungsi tujuan dengan menggunakan Persamaan 1. Setelah itu, menghitung nilai keanggotaan untuk setiap data pada tiap *centroid* dengan menggunakan Persamaan 2 dan menghitung nilai bobot pada setiap data dengan menggunakan Persamaan 3. Selanjutnya melihat nilai fungsi tujuan, apabila nilai fungsi tujuan berubah maka perhitungan terus dilakukan hingga nilai fungsi tujuan tidak berubah secara signifikan. Selanjutnya memasukkan setiap data pada tiap *cluster* berdasarkan nilai keanggotaan terbesar. Hasil dari *cluster* tersebut divalidasi dengan menggunakan nilai *Dunn Index* pada Persamaan 5 hingga mendapatkan hasil *cluster* terbaik.

III. HASIL DAN PEMBAHASAN

A. Analisis Deskriptif Data

Pada analisis deskriptif ini bertujuan memberikan ringkasan data secara umum tentang variabel yang digunakan. Berikut hasil analisis deskriptif pada data tingkat kemiskinan Kabupaten/Kota di Indonesia ditunjukkan pada Tabel 2.

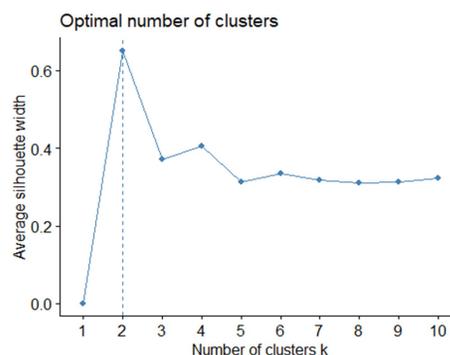
Tabel 2. Analisis Deskriptif

Variabel	Min	Max	Median	Mean	Simpangan Baku
X_1	2,28	42,03	9,82	11,68	7,27
X_2	0,17	13,90	1,37	1,97	1,84
X_3	0,02	7,59	0,31	0,53	0,73
X_4	264.666	1.099.019	469.719	489.964	116.214

Berdasarkan Tabel 2, dapat dilihat bahwa pada tiap variabel memiliki nilai minimum dan maksimum yang sangat timpang. Dimana nilai minimum pada variabel X_1 yaitu sebesar 2,28 persen yang berada di Kota Sawahlunto, sedangkan nilai maksimum pada variabel X_1 yaitu sebesar 42,03 persen yang berada di Kabupaten Intan Jaya. kabupaten/kota dengan X_2 dan X_3 terendah yaitu sebesar 0,17 dan 0,02 persen berada di Kota Banjar Baru, sedangkan kabupaten/kota dengan dengan X_2 dan X_3 tertinggi yaitu 13,90 dan 7,59 persen berada di Kabupaten Puncak Jaya dan Kabupaten Lanny Jaya. Pada variabel X_4 memiliki nilai terendah sebesar 264.666 rupiah per kapita per bulan berada di Kabupaten Buton Selatan, sedangkan Kabupaten/Kota dengan nilai tertinggi yaitu sebesar 1.099.019 rupiah per kapita per bulan berada di Kota Jayapura.

B. Analisis Cluster K-Harmonic Means (KHM)

Selanjutnya dibentuk jumlah kelompok atau *cluster*, jumlah *cluster* yang optimal ditentukan berdasarkan nilai *silhouette coefficient* tertinggi. Adapun nilai *silhouette coefficient* pada tiap jumlah *cluster* ditunjukkan pada Gambar 1.



Gambar 1. Nilai Silhouette Coefficient

Berdasarkan Gambar 1, dapat dilihat bahwa nilai *silhouette coefficient* tertinggi berada pada jumlah *cluster* 2. Dengan demikian pengelompokkan tingkat kemiskinan kabupaten/kota di Indonesia membentuk 2 *cluster* pada pendekatan KHM. Pendekatan KHM dihitung dengan menggunakan beberapa parameter agar mendapatkan hasil pengelompokkan yang optimal. Parameter yang digunakan pada pendekatan KHM yaitu $p = 2; 2,25$ dan $2,5$. Selanjutnya menghitung nilai fungsi tujuan dengan Persamaan 1 pada KHM ditunjukkan pada Tabel 3.

Tabel 3. Nilai Fungsi Tujuan Metode KHM

Parameter (p)	Nilai Fungsi	Iterasi
	Tujuan	
2	1760,857	31
2,25	2235,764	21
2,5	2863,212	17

Berdasarkan Tabel 3, didapatkan nilai fungsi tujuan dengan menggunakan $p = 2; 2,25$ dan $2,5$ serta banyak iterasi yang didapat. Selanjutnya menghitung nilai keanggotaan menggunakan Persamaan 2, nilai keanggotaan pada pendekatan KHM ini digunakan untuk menentukan data pengamatan masuk kedalam *cluster* dengan melihat besaran peluang data pengamatan masuk ke dalam *cluster*. Hasil nilai keanggotaan pada metode KHM dapat dilihat pada Tabel 4.

Tabel 4. Nilai Keanggotaan Metode KHM

Menggunakan $p = 2$		
Data ke-	c_1	c_2
1	0,278	0,722
2	0,601	0,398
⋮	⋮	⋮
514	0,505	0,495
Menggunakan $p = 2,25$		
Data ke-	c_1	c_2
1	0,055	0,945
2	0,166	0,834
⋮	⋮	⋮
514	0,434	0,566
Menggunakan $p = 2,5$		
Data ke-	c_1	c_2
1	0,015	0,985
2	0,049	0,950
⋮	⋮	⋮
514	0,358	0,642

Berdasarkan Tabel 4, didapat nilai keanggotaan dengan menggunakan $p = 2; 2,25$ dan $2,5$. Maka dihitung bobot dengan menggunakan Persamaan 3, hasil dari perhitungan bobot ditunjukkan pada Tabel 5.

Tabel 5. Nilai Bobot Metode KHM

Data ke-	Nilai Bobot Titik Data		
	$p = 2$	$p = 2,25$	$p = 2,5$
1	0,527	0,781	0,997
2	0,505	0,679	0,957
⋮	⋮	⋮	⋮
514	0,500	0,770	1,211

Berdasarkan Tabel 5, didapat hasil dari bobot dengan menggunakan $p = 2; 2,25$ dan $2,5$. Setelah didapatkan nilai keanggotaan dan nilai bobot pada setiap data, maka centroid baru dapat dihitung dengan menggunakan Persamaan 4. Setelah itu ditetapkan keanggotaan data pengamatan pada suatu *cluster* dengan melihat nilai tertinggi terhadap nilai keanggotaan ketika iterasi berhenti. Tabel 6 menampilkan hasil keanggotaan pada data pengamatan di tiap *cluster*.

Tabel 6. Hasil Penetapan Keanggotaan Data Metode KHM

Menggunakan $p = 2$		
Cluster	1	2
Jumlah Kab/Kota	73	441
Menggunakan $p = 2,25$		
Cluster	1	2
Jumlah Kab/Kota	54	460
Menggunakan $p = 2,5$		
Cluster	1	2
Jumlah Kab/Kota	45	469

Berdasarkan Tabel 6, hasil pengelompokkan dengan $p = 2$ pada data pengamatan adalah 73 kabupaten/kota yang bergabung dengan *cluster* 1 dan pada *cluster* 2 didapat 441 kabupaten/kota. Pada pengelompokkan dengan $p = 2,25$ didapatkan anggota pada *cluster* 1 adalah 54 kabupaten/kota dan *cluster* 2 beranggotakan 460 kabupaten/kota. Sedangkan hasil yang didapatkan dengan menggunakan $p = 2,5$ adalah 45 kabupaten/kota tergabung pada *cluster* 1 dan 469 kabupaten /kota tergabung pada *cluster* 2. Hasil dari *cluster* tersebut dibandingkan berdasarkan nilai *Dunn Index* yang dihitung dengan Persamaan 5. Perbandingan hasil validasi *cluster* tersebut dapat dilihat pada Tabel. 7

Tabel 7. Validasi Hasil *Cluster* KHM

Validasi	$p = 2$	$p = 2,25$	$p = 2,5$
Dunn Index	0,00717	0,03492	0,03165

Berdasarkan Tabel 7, nilai *Dunn Index* yang digunakan untuk memvalidasi hasil pengelompokkan tingkat kemiskinan kabupaten/kota di Indonesia menggunakan pendekatan *K-Harmonic Means* didapatkan hasil yang berbeda. Pengelompokkan dengan $p = 2,25$ memiliki Nilai *Dunn Index* tertinggi yaitu sebesar 0,03492. Oleh karena itu, diputuskan bahwa pengelompokkan yang optimal pada pendekatan *K-Harmonic Means* adalah dengan menggunakan $p = 2,25$. Karakteristik dari masing-masing *cluster* ditunjukkan dari nilai rata-rata *cluster* yang ditampilkan pada Tabel 8.

Tabel 8. Rata-rata *Cluster* Metode KHM dengan $p = 2,25$

Cluster	X_1	X_2	X_3	X_3	Kategori
c_1	2,107	2,186	1,993	0,606	Tinggi
c_2	-0,297	-0,292	-0,256	-0,079	Rendah

Berdasarkan Tabel 8, kabupaten/kota yang masuk kedalam *cluster* pertama memiliki nilai pada setiap variabelnya di atas rata-rata, dapat dilihat bahwa semua nilainya bernilai positif dan nilainya lebih tinggi dari pada *cluster* kedua. Sehingga kabupaten/kota yang masuk kedalam *cluster* pertama ini adalah kabupaten/kota yang memiliki tingkat kemiskinan yang tinggi, maka penanganannya menjadi prioritas pertama. Sedangkan kabupaten/kota yang masuk ke dalam *cluster* kedua ini memiliki nilai setiap variabelnya berada di bawah rata-rata, dapat dilihat pada Tabel 8 bahwa semua nilainya negatif dan nilai rata-rata pada *cluster* ini lebih kecil pada *cluster* pertama. Sehingga kabupaten/kota yang masuk ke dalam *cluster* kedua ini adalah kabupaten/kota dengan tingkat kemiskinan rendah, maka penanganannya menjadi prioritas kedua.

IV. KESIMPULAN

Hasil pengelompokkan tingkat kemiskinan kabupaten/kota di Indonesia dengan pendekatan *K-Harmonic Means* diperoleh pengelompokkan yang paling optimal dengan menggunakan $p = 2,25$. Pendekatan *K-Harmonic Means* dengan menggunakan $p = 2,25$ diperoleh jumlah *cluster* sebanyak 2 *cluster* dengan kategori yaitu *cluster* 1 dengan tingkat kemiskinan tinggi, dimana *cluster* 1 beranggotakan 54 kabupaten/kota, sedangkan *cluster* 2 dikategorikan dengan tingkat kemiskinan rendah yang beranggotakan 460 Kabupaten/Kota. Hal tersebut menunjukkan bahwa banyak dari kabupaten/kota yang memiliki tingkat kemiskinan yang rendah. Dengan diketahui tingkat kemiskinan dari masing-masing kabupaten/kota dapat menjadi acuan bagi pemerintah atau pembuat kebijakan dalam menanggulangi kemiskinan lebih mudah untuk mengalokasikan anggaran yang lebih tepat sasaran. Penelitian selanjutnya disarankan

untuk menggunakan analisis *cluster* yang lainnya serta dapat membandingkan analisis *cluster K-Harmonic Means* dengan metode *cluster* lainnya.

DAFTAR PUSTAKA

- Adhitya, B., Prabawa, A., & Kencana, H. (2022). Analisis Pengaruh Pendidikan, Kesehatan, Sanitasi dan Rata-Rata Jumlah Anggota Keluarga Per Rumah Tangga terhadap Kemiskinan di Indonesia. *Ekonomis: Journal of Economics and Business*, 6(1), 288-295.
- Arifin, J. (2020). Budaya kemiskinan dalam penanggulangan kemiskinan di Indonesia. *Sosio Informa: Kajian Permasalahan Sosial dan Usaha Kesejahteraan Sosial*, 6(2), 114-132.
- Badan Pusat Statistik (BPS). "Kemiskinan dan Ketimpangan". Publikasi BPS diakses dari <http://www.bps.go.id/> , tanggal 15 April 2022
- Christienova, S. I., Suparman, Y., & Zulhanif, Z. (2017). Clustering dengan Metode K-Means dan K-Harmonic Means: Pengelompokan Sister Village Daerah Sulit dalam Survei. In *E-Prosiding Seminar Nasional Statistika| Departemen Statistika FMIPA Universitas Padjadjaran* (Vol. 6, pp. 33-40).
- Gunaidi, I. G. A. (2019). Analisis Cluster Pada Pengelompokan Siswa Diktuk Bintara Polri TA. 2018/2019, SPN Singaraja – Polda Bali Menggunakan K-Means dan K-Harmonic Means. *Jurnal Ilmiah SINUS*, 17(2), 13. <https://doi.org/10.30646/sinus.v17i2.421>
- Hair, J. Black, W., Babin, B., & Anderson, R. (2009). *Multivariate Data Analysis*. 7th ed. Edinburgh: Pearson.
- Hamerly, G., & Elkan, C. (2002). Alternative to the k-means algorithm that find better clusterings. *International Conference on Information and Knowledge Management, Proceedings*, 600-607. <https://doi.org/10.1145/584792.584890>
- Hermawati, F. A. (2013). *Data Mining*. Yogyakarta: Penerbit Andi
- Hung, C. H., Chiou, H. M., & Yang, W. N. (2013). Candidate groups search for K-Harmonic Means data clustering. *Applied Mathematical Modelling*, 37(24), 10123-10128. <https://doi.org/10.1016/j.apm.2013.05.052>
- Putra, R. F. I., & Lisna, V. (2020). Segitiga Kemiskinan-Pertumbuhan-Ketimpangan (PGI TRIANGLE): Pembangunan Keuangan, Pembangunan Manusia, dan Ketimpangan Pendapatan Di Asia. *Jurnal Ekonomi dan Pembangunan*, 28(2), 77-89.
- Trianto. (2010). *Model Pembelajaran Terpadu, Konsep, Strategi dan Implementasinya dalam KTSP*. Jakarta: Bumi Aksara.
- Wierzchon, S. T., & Klopotek, M. A. (2018). *Modern Algorithms of Cluster Analysis*. Poland: Springer