

# Analisis Biplot Pada Pengelompokan Kecamatan Di Kabupaten Tasikmalaya Berdasarkan Indikator Kemiskinan

Annisa Siti Utami<sup>1\*</sup>, Anindya Apriliyanti Pravitasari<sup>2</sup>, and Irlandia Ginanjar<sup>3</sup>

<sup>1,2,3</sup>Departemen Statistika, Universitas Padjadjaran, Indonesia  
\*Corresponding author: annisa19001@mail.unpad.ac.id

Received: 24 January 2023

Revised: 4 August 2023

Accepted: 4 October 2023

**ABSTRAK** – Kemiskinan merupakan masalah sosial yang terus ada dalam kehidupan masyarakat menurut Nurwati, 2008. Oleh karena itu, permasalahan kemiskinan menjadi pusat perhatian bagi pemerintah Kabupaten Tasikmalaya. Dalam rencana pembangunan jangka panjang nasional (RPJPN) 2005-2025 masalah kemiskinan dilihat dalam kerangka multidimensi karenanya kemiskinan bukan hanya terkait ukuran pendapatan, melainkan menyangkut beberapa hal. Karena, kemiskinan bukan hanya terkait ukuran pendapatan melainkan menyangkut beberapa hal. Dalam rencana pembangunan jangka menengah daerah (RPJMD) Kabupaten Tasikmalaya, target capaian angka kemiskinan pada tahun 2021 sebesar 10,23%. Berdasarkan publikasi BPS terdapat sebanyak 10,75% penduduk Kabupaten Tasikmalaya yang dikategorikan sebagai penduduk miskin, artinya target pemerintah Kabupaten Tasikmalaya tidak tercapai. Sehingga perlu adanya upaya penanggulangan terhadap masalah kemiskinan. Penelitian ini bertujuan untuk mengelompokkan kecamatan di Kabupaten Tasikmalaya berdasarkan kemiripan indikator kemiskinan yang dimiliki oleh setiap kecamatan menggunakan analisis biplot. Data yang digunakan adalah data indikator kemiskinan 39 kecamatan di Kabupaten Tasikmalaya pada tahun 2021. Dari hasil penelitian didapatkan ukuran keragaman yang dapat digambarkan sebesar 97% artinya plot yang terbentuk dapat menggambarkan informasi data sebenarnya dengan baik. Selain itu, diperoleh tiga *cluster* yang memiliki kemiripan dalam indikator kemiskinan. *Cluster 1* berisi kecamatan yang memiliki indikator berupa tingginya rasio murid terhadap sekolah pada SMA/SMK/MA. *Cluster 2* berisi kecamatan yang memiliki indikator yang menengah hingga rendah pada semua variabel kecuali variabel rasio murid terhadap sekolah pada SMP/MTs dan rasio murid terhadap sekolah pada SMA/SMK/MA. Sedangkan *Cluster 3* berisi kecamatan yang memiliki indikator berupa tingginya rasio murid terhadap sekolah pada SMP/MTs.

**Kata kunci** – kemiskinan, *Sustainable Development Goals* (SDGs), analisis biplot

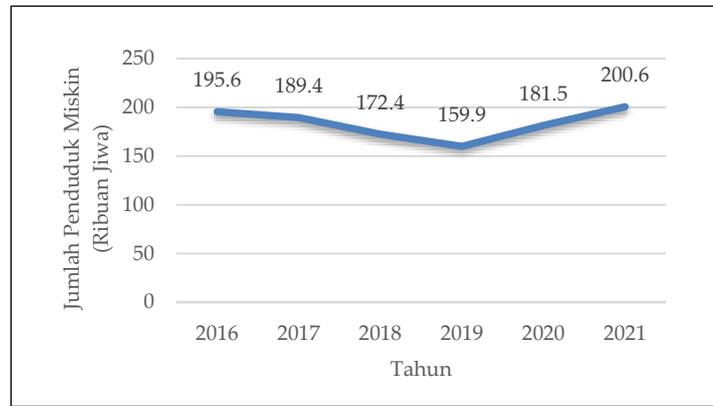
**ABSTRACT** – Poverty is a social problem that continues to exist in people's lives according to Nurwati, 2008. Therefore, the problem of poverty is the center of attention of the Tasikmalaya Regency government. In the National Long-Term Development Plan (RPJPN) 2005-2025 the problem of poverty is seen in a multidimensional framework, therefore poverty is not only related to income measurement, but related to several things. This is because poverty is not only related to the size of income but involves several things. In the Tasikmalaya Regency Regional Medium-Term Development Plan (RPJMD), the target for achieving the poverty rate in 2021 is 10.23%. Based on BPS publications, there are 10.75% of the population of Tasikmalaya Regency who are categorized as poor, meaning that the Tasikmalaya Regency government's target has not been achieved. So it is necessary to make efforts to overcome the problem of poverty. This study aims to group sub-districts in Tasikmalaya Regency based on the similarity of poverty indicators owned by each sub-district by using biplot analysis. The data used is poverty indicator data for 39 sub-districts in Tasikmalaya Regency in 2021. From the research results it is known that the amount of variation that can be described is 97%, meaning that the plots formed can best describe actual conditions. data information. In addition, three clusters have the same poverty indicators. Cluster 1 contains sub districts that have an indicator in the form of a high student to school ratio in SMA/SMK/MA. Cluster 2 contains sub districts that have moderate to low indicators on all variables except the ratio of SMP/MTs students and the ratio of SMA/SMK/MA students. Meanwhile, Cluster 3 consists of sub-districts that have an indicator in the form of a high ratio of SMP/MTs students.

**Keywords** – poverty, *Sustainable Development Goals* (SDGs), biplot analysis

## I. PENDAHULUAN

Kemiskinan merupakan masalah sosial yang terus ada dalam kehidupan masyarakat menurut [1]. Oleh karena itu, permasalahan kemiskinan menjadi pusat perhatian untuk pemerintah pusat maupun pemerintah kabupaten dan kota yang ada di Indonesia, tidak terkecuali dengan pemerintah Kabupaten Tasikmalaya. Melalui Instruksi Presiden Republik Indonesia (Inpres) Nomor 4 Tahun 2022 tentang Percepatan Penghapusan Kemiskinan Ekstrem. Pemerintah melakukan percepatan dalam penanggulangan masalah kemiskinan ekstrem dengan target 0% pada tahun 2024.

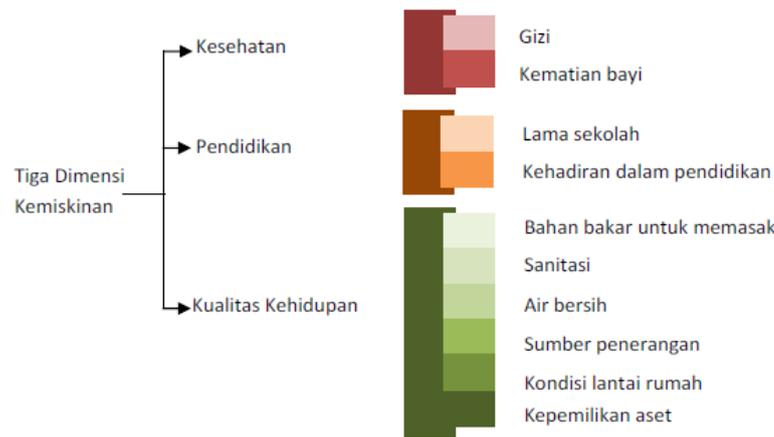
Dalam rencana pembangunan jangka Panjang nasional (RPJPN) 2005-2025, masalah kemiskinan dilihat dalam kerangka multidimensi, karenanya kemiskinan bukan hanya terkait ukuran pendapatan, melainkan menyangkut beberapa hal, antara lain: (i) kerentanan dan kerawanan orang atau masyarakat untuk menjadi miskin; (ii) menyangkut ada/tidak adanya pemenuhan hak dasar warga dan ada/tidak adanya perbedaan perlakuan seseorang atau kelompok masyarakat dalam menjalani kehidupan secara bermartabat. Hal tersebut diyakini agar melihat kemiskinan dari berbagai dimensi dan memandang penyebab kemiskinan dari berbagai sisi.



Gambar 1 Jumlah Penduduk Miskin di Kabupaten Tasikmalaya 2016-2021

Dalam rencana pembangunan jangka menengah daerah (RPJMD) Kabupaten Tasikmalaya, target capaian angka kemiskinan pada tahun 2021 sebesar 10,23%. Dapat dilihat pada Gambar 1, pada tahun 2021 tercatat ada sebanyak 200,6 ribu atau 10,75% penduduk Kabupaten Tasikmalaya yang dikategorikan sebagai penduduk miskin [2]. Artinya target pemerintah untuk angka kemiskinan tidak tercapai. Sehingga perlu upaya penanggulangan masalah kemiskinan.

United National Development Program (UNDP) dan Oxford Poverty and Human Development Initiative (OPHI) mengembangkan Global Multidimensional Poverty Index (MPI) yang merupakan ukuran internasional dari kemiskinan multidimensi. MPI tidak hanya memperhatikan tentang pendapatan untuk memahami tentang bagaimana kemiskinan itu sendiri terjadi. Melainkan bagaimana kemiskinan terjadi dalam keadaan di tiga dimensi yaitu Kesehatan, Pendidikan, dan standar hidup [3]. Sehingga kemiskinan itu sendiri bisa dilihat dari pengeluaran per kapita yang menunjukkan garis kemiskinan akan tetapi secara luas kemiskinan tidak hanya tentang pendapatan atau pengeluaran melainkan mampu diperhatikan dari berbagai dimensi seperti Kesehatan, Pendidikan, serta standar hidup secara individu maupun masyarakat suatu daerah. Sehingga dalam penelitian ini melibatkan beberapa variabel yang mampu mewakili ketiga dimensi tersebut.



Gambar 2 Dimensi dan Indikator setiap Dimensi dalam MPI

Setiap kecamatan mempunyai nilai indikator yang berbeda-beda, sehingga program pengentasan kemiskinan untuk setiap kecamatan tidak sama. Perbedaan nilai tersebut, menunjukkan bahwa identifikasi homogenitas indikator kemiskinan untuk setiap kecamatan membutuhkan pengelompokan. Oleh karena itu, dibutuhkan pengelompokan yang tepat dalam mengakomodir wilayah pada tiap kecamatan yang memiliki persamaan indikator pada wilayah tersebut. Salah satu metode yang dapat digunakan adalah analisis biplot. Analisis biplot merupakan Teknik statistik dimensi ganda dengan menyajikan secara visual dan simultan sejumlah objek pengamatan dan variabel dalam suatu grafik.

Tujuan dari penelitian ini adalah menentukan kelompok berdasarkan kemiripan nilai indikator yang dimiliki oleh setiap kecamatan agar bermanfaat sebagai masukan bagi penentuan kebijakan penanganan kemiskinan untuk meningkatkan kesejahteraan rakyat di Kabupaten Tasikmalaya. Manfaat dari penelitian ini adalah memberikan referensi untuk Badan Perencanaan Pembangunan, Penelitian dan Pengembangan Daerah (Bappelitbangda) Kabupaten Tasikmalaya dalam menentukan kecamatan prioritas yang memerlukan bantuan berdasarkan indikator kemiskinan.

**II. LITERATURE REVIEW**

**A. Analisis Biplot**

Analisis biplot merupakan salah satu Teknik peubah ganda yang menyajikan plot pengamatan  $n$  dan variabel  $p$  secara bersamaan dalam bidang dua dimensi. Penyajian plot pengamatan  $n$  dan variabel  $p$  secara bersamaan dapat memberikan tambahan informasi yang lebih baik tentang hubungan antara variabel dan pengamatan [4]. Terdapat 4 hal penting yang bisa didapatkan dari tampilan biplot. yaitu kemiripan antar objek, keragaman peubah, korelasi antar peubah dan nilai relatif yang dimiliki suatu objek [5].

Keberadaan asosiasi adalah dasar yang valid untuk pengelompokan objek berdasarkan nilai sumbu utama, jadi jika sumbu utama digunakan dalam PCA sebanyak  $m$ , maka objek dapat diklasifikasikan menjadi  $2^m$ . Adanya hubungan yang besar antara sumbu utama dan variabel awal sebagai dasar yang valid untuk menggunakan nilai sumbu utama sebagai ukuran untuk pengelompokan objek. Objek dapat dikelompokkan berdasarkan pada karakteristiknya (diwakili oleh sumbu utama) dalam dua kelompok, dimana objek dengan sumbu utama bernilai positif maka nilainya akan berada di atas rata-rata, sedangkan jika objek dengan sumbu utama bernilai negatif maka nilainya akan berada dibawah rata-rata, sedangkan jika  $m$  sumbu utama digunakan dalam PCA untuk  $m = 1, 2, \dots, r$ , maka objek dapat diklasifikasikan menjadi  $2^m$  kelompok. Pengelompokan objek berdasarkan nilai PCA Biplot lebih lengkap dituliskan dalam [6].

**B. Penguraian Singular Value Decomposition (SVD)**

*Singular Value Decomposition (SVD)* bertujuan untuk menguraikan matriks  $X$  berukuran  $n \times p$  dimana  $n$  adalah banyaknya objek pengamatan dan  $p$  adalah banyaknya variabel, menjadi 3 buah matriks. Jika dimisalkan sebuah matriks data berpangkat dengan ukuran dimana jumlah pengamatan dan jumlah variabel dikoreksi terhadap nilai rataannya, maka dapat diuraikan menjadi [7]:

$$X_{n \times p} = U_{n \times r} L_{r \times r} A'_{r \times p} \tag{1}$$

dimana

- $X_{n \times p}$  : matriks data berukuran  $(n \times p)$  yang telah dikoreksi terhadap rataannya
- $U_{n \times r}$  : matriks berukuran  $(n \times r)$  yang kolomnya disebut vektor singular kolom yang merupakan landasan orthonormal kolom-kolom matriks dalam ruang dimensi  $n$
- $L_{r \times r}$  : matriks berukuran  $(r \times r)$  dengan unsur diagonal utamanya adalah nilai singular matriks  $X$  yaitu akar kuadrat dari nilai eigen matriks  $X'X$
- $A'_{r \times p}$  : matriks berukuran  $(p \times r)$  yang kolomnya merupakan vektor eigen dari matriks  $X'X$  dimana kolom tersebut disebut sebagai vektor singular baris yang merupakan landasan orthonormal kolom-kolom matriks  $X$  dalam ruang dimensi  $p$

**C. Faktorisasi Matriks**

Analisis biplot dibangun menggunakan matriks data, dimana matriks data tersebut berupa variabel yang diwakili oleh masing-masing kolom dan objek penelitian yang diwakili oleh masing-masing baris. Matriks data tersebut dapat dikatakan matriks  $X$  yang memiliki sebanyak  $p$  variabel penelitian dan juga sebanyak  $n$  objek yang diteliti. Kemudian dapat dibentuk matriks  $G$  dan  $H$  dari matriks  $X$  tersebut. Jika didefinisikan  $G = UL^\alpha$  berukuran  $n \times r$  dan  $H' = L^{1-\alpha} A'$  berukuran  $r \times p$  dimana  $r$  adalah rank matriks data  $X$  dan  $0 \leq \alpha \leq 1$ .

$$X = UL^\alpha L^{1-\alpha} A' = GH' \tag{2}$$

Sehingga unsurke- $(i,j)$  dari matriks  $X$  dapat dirumuskan sebagai berikut :

$$X_{ij} = g'_i h_j \tag{3}$$

dimana:

- $g'_i = 1, 2, 3, \dots, n$  yaitu baris dari matriks  $G$
- $h_j = 1, 2, 3, \dots, p$  yaitu baris dari matriks  $H$

Vektor pengaruh baris  $g_i$  dan vector pengaruh kolom  $h_j$  dapat digambarkan dengan pasti dalam ruang berdimensi dua jika  $X$  berpangkat dua. Karena jika  $X$  mempunyai rank dua ( $r = 2$ ) dan  $g'_i$  serta  $h'_j$  keduanya memiliki  $r$  buah elemen, maka  $G$  berukuran  $n \times 2$  dan  $H$  berukuran  $p \times 2$ . Pemilihan nilai  $\alpha$  bersifat sembarang dengan syarat  $0 \leq \alpha \leq 1$ . Menurut Jolliffe pengambilan nilai ekstrim  $\alpha = 0$  dan  $\alpha = 1$  akan berguna untuk mempermudah interpretasi hasil biplot. Apabila nilai  $\alpha = 0$  maka  $G = U$  dan  $H = AL$  sehingga diperoleh persamaan :

$$X'X = HH' = (n - 1)S \tag{4}$$

Hasil kali  $h'_j h_k$  dimana  $h_j$  dan  $h_k$  ialah nilai pada baris ke- $j$  dan baris ke- $k$  dari matriks  $\mathbf{H}$  dengan  $j=k=1,2,\dots,p$  akan sama dengan  $(n - 1)$  dikali  $S_{jk}$  dan  $h'_j h_k$  menggambarkan keragaman peubah ke- $k$  yang dilihat dari nilai  $\|h_j\|$  yang merupakan panjang vektor dari  $h_j$ . Sementara korelasi antara peubah ke- $j$  dan baris ke- $k$  ditunjukkan oleh kosinus sudut antara vektor  $h_j$  dan  $h_k$  dengan persamaan sebagai berikut:

$$\cos\theta = \frac{h'_j h_k}{(h'_j h_j)(h'_k h_k)} \tag{5}$$

**D. Ukuran Keragaman Biplot**

Data yang digunakan dalam analisis biplot direduksi ke dalam dimensi yang lebih rendah. Sebagai ukuran dari pendekatan ini, perlu menghitung total inersia. Total inersia dapat menunjukkan persentase kategori atau informasi yang hilang. Total inersia dapat diperoleh dengan notasi berikut [8] :

$$\text{trace}(\mathbf{F}^T \mathbf{F}) = \text{trace}(\mathbf{\Lambda})$$

$$\tau_d = \left( \frac{\lambda_1 + \lambda_2 + \dots + \lambda_d}{\sum_{k=1}^r \lambda_k} \right)$$

Cakupan varians untuk setiap dimensi dinotasikan sebagai berikut [8] :

$$\tau_d = \frac{\sum_{k=1}^d \lambda_k}{\sum_{k=1}^r \lambda_k} \tag{6}$$

dimana :

$\lambda_k$  : nilai eigen ke- $k$  ( $k = 1,2, \dots, r$ )

$\tau_d$  : varians tercakup

Apabila mendekati nilai satu, maka biplot memberikan penyajian yang semakin baik mengenai informasi dari data sebenarnya.

**E. Analisis Cluster**

Hasil dari analisis biplot berupa peta persepsi yang kurang menunjukkan pengelompokkan secara tegas, sehingga koordinat dari biplot akan dilakukan analisis cluster agar dapat dihasilkan kelompok kecamatan berdasarkan karakteristik kemiskinan yang dimiliki. Analisis cluster adalah teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya tanpa menganalisis hubungan antara variabel independen dan dependen. Analisis cluster mengklasifikasi objek sehingga setiap objek yang memiliki sifat yang mirip akan mengelompok ke dalam satu cluster (kelompok) yang sama. Untuk mendapatkan kelompok yang sehomogen mungkin, maka yang digunakan dasar untuk mengelompokkan adalah kesamaan skor nilai yang dianalisis. Semakin kecil besaran jarak suatu individu terhadap individu lain, maka semakin besar kemiripan individu tersebut.

Dalam banyak kasus kesamaan didasarkan pada beberapa ukuran jarak. Jarak yang menjauh menggambarkan ketidakmiripan antar objek yang diteliti. Pada umumnya analisis cluster dibagi menjadi dua metode yaitu hierarchical cluster dan non hierarchical cluster [9].

**Hierarchical Cluster Analysis**

Metode hierarki merupakan teknik pengelompokan dimana jumlah cluster belum diketahui. Metode hierarki terbagi menjadi dua, yaitu dengan penggabungan (agglomerative) dan pemisahan (divise). Metode hierarki penggabungan, yakni pada awal pengelompokan setiap objek pengamatan dianggap sebagai cluster yang berbeda, kemudian secara bertahap objek-objek yang memiliki kemiripan dikelompokkan ke dalam cluster yang sama hingga pada akhirnya semua objek berada dalam satu cluster yang sama. Sedangkan metode hierarki pemisahan memiliki langkah pengerjaan yang berlawanan dengan metode hierarki penggabungan. Metode hierarki pemisahan, yaitu semua objek dianggap berasal dari satu cluster, kemudian dilihat ketidakmiripan antar objek. Objek yang tidak mirip akan dikeluarkan dari cluster dan membentuk cluster sendiri. Tahapan ini dilakukan sampai pada akhirnya semua cluster beranggotakan satu objek [10]. Pada Hierarchical Clustering data dikelompokkan melalui suatu bagan yang berupa hirarki, dimana terdapat penggabungan dua grup yang terdekat di setiap iterasinya ataupun pembagian dari seluruh set data kedalam cluster [11].

**Non-Hierarchical Cluster Analysis**

Pada metode ini dimulai dengan menentukan terlebih dahulu jumlah cluster yang akan dibuat, barulah setelah itu dilakukan analisis cluster tanpa melalui tahapan tingkatan hierarki seperti metode sebelumnya. Pada metode ini anggota cluster dapat berpindah ke cluster lain dalam prosesnya hingga tercapai homogenitas maksimal di dalam cluster [12]. Metode non-hirarki yang paling banyak digunakan adalah metode K-means. Pada non-hierarchical cluster analysis terbagi menjadi beberapa metode, diantaranya :

1. K-means merupakan metode yang bertujuan untuk mempartisi observasi menjadi cluster dimana setiap observasi termasuk dalam cluster dengan mean terdekat. Pada metode ini nilai centroid sebagai pusat cluster yang mewakili cluster itu sendiri [11]. Pengelompokkan dilakukan dengan meminimalkan jumlah kuadrat jarak antara data dan centroid cluster yang sesuai. Nilai rata-rata pada cluster (centroid) kemudian dihitung secara berulang hingga pengelompokkan berhenti saat tidak ada lagi objek yang berpindah cluster.
2. K-Medoids Clustering adalah metode pengelompokkan beberapa objek dalam suatu data ke dalam  $k$  buah cluster dengan menggunakan medoid sebagai pusat dari cluster tersebut. Penggunaan metode K-Medoids dianggap lebih robust terhadap outlier jika dibandingkan dengan algoritma yang lain serta lebih dapat digunakan untuk data yang lebih besar menurut [13] dalam [14].

**F. Penentuan Jumlah Cluster Optimum dengan Metode Elbow**

Pada metode *k-medoids clustering* terlebih dahulu harus menentukan jumlah *cluster* yang akan digunakan pada analisis. Pemilihan jumlah *cluster* sebaiknya tidak sembarang agar *cluster* yang dihasilkan optimal. Dengan menggunakan metode *Elbow* kita dapat menentukan jumlah cluster yang optimal. Metode *Elbow* diawali dengan menentukan nilai awal  $k$ , lalu menghitung nilai *Sum of Square Error* (SSE) dari masing-masing nilai  $k$  tersebut.

$$SSE = \sum_{k=1}^K \sum_{x_i \in S_k} \|x_i - C_k\|^2 \tag{7}$$

dimana :

- $K$  : jumlah *cluster*
- $x_i$  : data ke- $i$
- $C_k$  : centroid *cluster*

Pada metode *Elbow* nilai *cluster* terbaik yang akan diambil dari nilai SSE yang mengalami penurunan yang signifikan dan berbentuk siku saat digambarkan dalam bentuk grafik [15].

**G. Nilai Medoids**

Dalam metode *k-medoids clustering* pusat *cluster* yang berperan dalam pengelompokan merupakan satu objek actual untuk mewakili *cluster*, dimana sebagai medoids yang representative terhadap *cluster*-nya. Partisi objek dilakukan berdasarkan prinsip meminimalkan total *error* antara setiap objek dengan medoids yang sesuai. Dengan kriteria kesalahan mutlak yang digunakan, didefinisikan sebagai berikut [11]:

$$E = \sum_{i=1}^k \sum_{p \in C_i} dist(p, o_i) \tag{8}$$

dimana:

- $E$  : jumlah kesalahan mutlak untuk semua objek  $p$  dalam data
- $o_i$  : Medoids atau objek representative dari  $C_i$
- $C_i$  : *Cluster* ke- $i$

**III. METODOLOGI**

**A. Data Penelitian**

Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari Badan Perencanaan Pembangunan, Penelitian dan Pengembangan Daerah (Bappelitbangda) Kabupaten Tasikmalaya dan Publikasi Badan Pusat Statistik Kabupaten Tasikmalaya dengan judul Kabupaten Tasikmalaya Dalam Angkatan 2022. Data penelitian ini menggunakan data nilai indikator kemiskinan yang dimiliki 39 kecamatan di Kabupaten Tasikmalaya pada tahun 2021, yang terdiri dari :

- $X_1$  : Persentase Penduduk yang Mengalami Gizi Buruk
- $X_2$  : Persentase Kematian Bayi
- $X_3$  : Persentase Penduduk dengan Usia Sekolah hingga Pra Sekolah
- $X_4$  : Persentase Penduduk dengan Usia Sekolah hingga SD/Sederajat
- $X_5$  : Persentase Penduduk dengan Usia Sekolah hingga SMP/Sederajat
- $X_6$  : Persentase Penduduk dengan Usia Sekolah hingga SMA/Sederajat
- $X_7$  : Persentase Penduduk dengan Usia Sekolah hingga Perguruan Tinggi
- $X_8$  : Rasio Murid terhadap Sekolah pada SD/MI
- $X_9$  : Rasio Murid terhadap Sekolah pada SMP/MTs
- $X_{10}$  : Rasio Murid terhadap Sekolah pada SMA/SMK/MA
- $X_{11}$  : Persentase Rumah yang Memiliki Akses Sanitasi Layak
- $X_{12}$  : Persentase Rumah yang Memiliki Akses Air Bersih

**B. Prosedur Penelitian**

Penelitian ini akan dilakukan melalui dua tahapan analisis, yaitu analisis biplot dan analisis k-medoids clustering.

**Analisis Biplot**

1. Menyusun data pengamatan dalam bentuk matriks  $X_{n \times p}$
2. Menghitung matriks  $X'X$
3. Menghitung nilai eigen dari matriks  $X'X$  dan pilih 2 nilai eigen terbesar
4. Menentukan matriks  $U$ ,  $L$ , dan  $A$  dari transformasi matriks dengan SVD

$$X_{n \times p} = U_{n \times r} L_{r \times r} A'_{r \times p} \tag{9}$$

dimana :

$$L = [\sqrt{\lambda_1} \dots 0 \dots 0 \dots \sqrt{\lambda_i}] \tag{10}$$

$$A = [a_1 \dots a_r] \tag{11}$$

$$U = \frac{1}{\sqrt{\lambda_i}} X a_i \tag{12}$$

5. Menghitung ukuran kelayakan biplot dari dua nilai eigen terbesar bila nilainya cukup besar ( $\geq 70\%$ ) maka pendekatan biplot dapat digunakan untuk memberikan penyajian visual bagi matriks data  $X$

$$\tau_2 = \frac{(\lambda_1 + \lambda_2)}{\sum_{k=1}^r \lambda_k} \tag{13}$$

dimana :

$\lambda_1$  = nilai eigen terbesar ke-1

$\lambda_2$  = nilai eigen terbesar ke-2

$\lambda_k$  = nilai eigen ke- $k$  ( $k = 1, 2, \dots, r$ )

6. Membuat matriks baris  $G$  dan kolom  $H'$  yang dibentuk dari SVD matriks  $X$

$$G = UL^{\alpha} \tag{14}$$

$$H' = L^{1-\alpha} A' \tag{15}$$

7. Membuat gambar biplot dengan menggunakan matriks  $G$  dan  $H$  sebagai koordinat
8. Mengelompokkan titik objek pada biplot berdasarkan kedekatan jarak dengan menggunakan k-medoids clustering.

**Analisis K-Medoids Clustering**

1. Data yang digunakan dalam analisis K-Medoids Clustering adalah data koordinat dari hasil analisis biplot yang akan dilakukan pengelompokkan
2. Menentukan banyak  $K$  (cluster) optimum menggunakan dengan metode Elbow
3. Menentukan nilai medoids sebanyak  $K$  buah
4. Menghitung jarak antara objek lain dengan medoids
5. Memilih nilai objek yang akan dijadikan sebagai medoids baru dan menghitung Kembali jarak objek dengan medoids yang baru
6. Menghitung total simpangan (S) dengan menghitung nilai total kesalahan mutlak (E) dengan medoid baru dikurangi dengan nilai total kesalahan mutlak (E) dengan medoid sebelumnya

**IV. HASIL DAN PEMBAHASAN**

Data yang digunakan dalam penelitian ini adalah data indikator kemiskinan yang dimiliki oleh 39 kecamatan di Kabupaten Tasikmalaya. Kemudian data indikator kemiskinan dari setiap kecamatan disusun dalam matriks  $X$ , sebagai berikut :

$$X_{39 \times 12} = [0.020 \ 0.521 \ 0.002 \ 0.135 \ \dots \ 13.66 \ \dots \ 15.99 \ \vdots \ 0.000 \ 0.000 \ \vdots \ \dots \ 14.52]$$

Setelah data dikoreksi terhadap rata-rata, kemudian matriks tersebut diuraikan menjadi tiga buah matriks menggunakan Persamaan (9). Dari hasil *Singular Value Decomposition* akan diperoleh matriks  $U$ , matriks  $L$ , dan matriks  $A$  sebagai berikut :

$$U_{39 \times 12} = [0.021957761 \quad -0.058120277 \quad -0.042418891 \quad -0.073159682 \quad \dots \quad -0.13367338 \quad \dots \quad 0.02266241 \quad \vdots \\ \vdots \quad 0.047948239 \quad 0.019917252 \quad \vdots \quad \vdots \quad \dots \quad 0.10761478 \quad ]$$

$$L_{12 \times 12} = [965.5993 \quad \dots \quad 0 \quad \vdots \quad \vdots \quad 0 \quad \dots \quad 0.1155653 \quad ]$$

$$A_{12 \times 12} = [-2.062747e - 05 \quad \dots \quad -0.9996364785 \quad \vdots \quad \vdots \quad -6.652818e - 04 \quad \dots \quad 0.0032453382 \quad ]$$

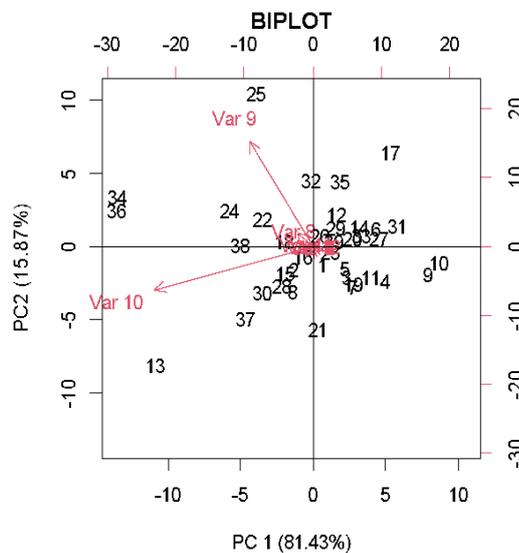
Dari matriks  $L$  yang unsur diagonalnya dikuadratkan akan menghasilkan nilai eigen yang dapat dibentuk menjadi matriks  $\lambda$  sebagai berikut :

$$\lambda = [9.323820e + 05 \quad \vdots \quad 1.335534e - 02 \quad ]$$

Selanjutnya akan dilakukan analisis biplot dan diperoleh matriks  $G$  dan  $H$  menggunakan persamaan (14) dan (15) yang selanjutnya diambil dua kolom pertama dari matriks  $G$  dan  $H$  yang dinotasikan sebagai matriks  $G^2$  dan  $H^2$ . Matriks  $G^2$  merupakan titik-titik objek pengamatan yaitu 39 kecamatan di Kabupaten Tasikmalaya, sedangkan matriks  $H^2$  merupakan titik-titik variabel pengamatan yang terdiri dari 8 indikator kemiskinan di desa/kelurahan. Berikut adalah matriks  $G^2$  dan  $H^2$  :

$$G^2_{39 \times 2} = [0.6823175 \quad -1.1999975 \quad \vdots \quad \vdots \quad 1.4899481 \quad 0.4112274 \quad ] \text{ dan } H^2_{12 \times 2} = [0.000640980 \quad - \\ 0.0007355621 \quad \vdots \quad \vdots \quad -0.020673030 \quad 0.1452176279 \quad ]$$

Kemudian dihasilkan ukuran keragaman dari (5) dengan jumlah dimensi sebanyak 2 dimensi menghasilkan keragaman sebesar 97.3% artinya biplot mampu menggambarkan informasi dari data sebenarnya maka dikatakan baik. Karena dengan 2 dimensi sudah mampu menggambarkan data sebenarnya dengan baik maka dihasilkan sebuah biplot seperti pada gambar berikut dengan bantuan software R.



Gambar 3 Biplot hasil pengelompokan

Berdasarkan Gambar 3. Didapatkan hasil pengelompokan kecamatan di Kabupaten Tasikmalaya berdasarkan indikator kemiskinan menggunakan analisis biplot, diperoleh informasi sebagai berikut :

**Kelompok Objek Penelitian**

Dari biplot diatas, dapat diperoleh informasi dimana terdapat 4 kelompok yang memiliki kemiripan indikator yang dapat dilihat dari kedekatan titik dalam biplot sebagai berikut :

- Kelompok 1 terdiri dari 13 kecamatan yaitu Kecamatan Cibalong, Sodonghilir, Salawu, Sukaraja, Cineam, Sukarame, Cigalontang, Padakembang, Sukaratu, Cisayong, Sukahening, Jamanis, dan Sukaresik.
- Kelompok 2 terdiri dari 7 kecamatan yaitu Kecamatan Salopa, Manonjaya, Singaparna, Mangunreja, Rajapolah, Ciawi, dan Pagerageung.
- Kelompok 3 terdiri dari 8 kecamatan yaitu Kecamatan Karangnunggal, Bantarkalong, Taraju, Puspahiang, Tanjungjaya, Leuwisari, Sariwangi, dan Kadipaten.
- Kelompok 4 terdiri dari 11 kecamatan yaitu Kecamatan Cipatujah, Cikalong, Pancatengah, Cikatomas, Parungponteng, Bojongasih, Culamega, Bojonggambir, Jatiwaras, Karangjaya, Gunungtanjung.

**Keragaman Variabel**

Berdasarkan hasil pada Tabel 1. dapat diperoleh informasi bahwa variabel rasio murid terhadap sekolah pada SMA/SMK/MA (X10) memiliki keragaman yang paling besar karena panjang vector variabelnya merupakan vector variabel yang paling panjang sementara variabel persentase penduduk yang mengalami gizi buruk (X1) memiliki keragaman yang paling kecil karena memiliki panjang vector variabel yang paling pendek.

**Tabel 1** Panjang Vektor Indikator Kemiskinan Kecamatan

Variabel	Panjang
X <sub>1</sub>	9.756572e - 04
X <sub>2</sub>	1.584997e - 02
X <sub>3</sub>	3.943814e - 02
X <sub>4</sub>	2.682258e - 02
X <sub>5</sub>	1.400070e - 02
X <sub>6</sub>	3.902652e - 02
X <sub>7</sub>	1.075371e - 02
X <sub>8</sub>	3.659477e + 00
X <sub>9</sub>	2.216158e + 01
X <sub>10</sub>	2.976969e + 01
X <sub>11</sub>	1.049663e + 00
X <sub>12</sub>	1.466817e - 01

**Korelasi Antar Variabel**

Dari hasil biplot diperoleh informasi berupa korelasi antar variabel, dalam penelitian ini korelasi antar indikator kemiskinan kecamatan. Berdasarkan persamaan (1.10), didapatkan hasil berupa matriks **cos cos θ** berukuran (12 × 12) yang berisi besar sudut yang dihasilkan antar vektor variabel indikator kemiskinan kecamatan berikut

**Tabel 2** Matriks *cos cos θ*

	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>
X <sub>1</sub>	1					
X <sub>2</sub>	-0.6639291	1				
X <sub>3</sub>	-0.5790073	-0.2252746	1			
X <sub>4</sub>	-0.9998562	0.6765152	0.56509727	1		
X <sub>5</sub>	-0.2799606	-0.5320183	0.94481804	0.2636398	1	
X <sub>6</sub>	-0.9816750	0.5092606	0.72376711	0.9783021	0.4577727	1
X <sub>7</sub>	-0.9940647	0.5786351	0.66427038	0.9920767	0.3827394	0.9965799
X <sub>8</sub>	-0.9921603	0.5652711	0.67636009	0.9898983	0.3977398	0.9977939
X <sub>9</sub>	-0.9854813	0.7812534	0.43217219	0.9882189	0.1129013	0.9350679
X <sub>10</sub>	-0.4352739	-0.3842479	0.98605990	0.4199434	0.9861560	0.5988608
X <sub>11</sub>	-0.9410044	0.3717105	0.82074883	0.9351304	0.5883065	0.9882459
X <sub>12</sub>	-0.8389816	0.9639445	0.04211094	0.8480892	-0.2875178	0.7199107

**Tabel 3** Matriks *cos cos θ* (Lanjutan)

	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	X <sub>11</sub>	X <sub>12</sub>
X <sub>1</sub>						
X <sub>2</sub>						
X <sub>3</sub>						
X <sub>4</sub>						
X <sub>5</sub>						
X <sub>6</sub>						
X <sub>7</sub>	1					
X <sub>8</sub>	0.9998673	1				
X <sub>9</sub>	0.9611612	0.9565373	1			
X <sub>10</sub>	0.5306346	0.5443730	0.2760980	1		
X <sub>11</sub>	0.9722335	0.9759169	0.8698883	0.7142503	1	
X <sub>12</sub>	0.7748024	0.7643999	0.9191903	-0.1247192	0.6053449	1

Berdasarkan hasil tersebut, diperoleh informasi mengenai korelasi antar variabel indikator kemiskinan sebagai berikut :

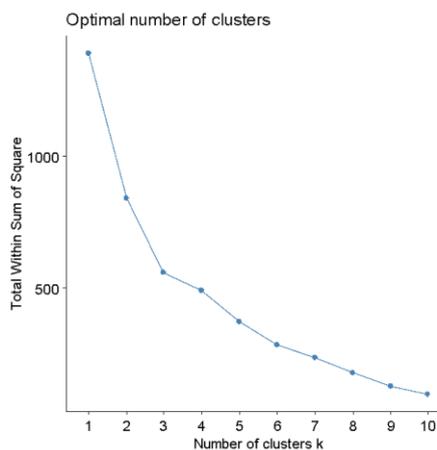
- Indikator kemiskinan yang memiliki korelasi tinggi adalah variabel persentase penduduk yang mengalami gizi buruk dengan variabel rasio murid terhadap sekolah pada SD/MI dan pasangan variabel lainnya yang memiliki nilai cosinus  $> 0.5$
- Indikator kemiskinan yang memiliki korelasi sedang adalah variabel persentase penduduk dengan usia sekolah hingga Pra Sekolah dengan variabel rasio murid terhadap sekolah pada SMP/MTs dan pasangan lainnya yang memiliki nilai cosinus diantara 0.3 hingga 0.5
- Indikator kemiskinan yang memiliki korelasi negatif sedang adalah variabel persentase penduduk yang mengalami gizi buruk dengan variabel rasio murid terhadap sekolah pada SMA/SMK/MA dan pasangan lainnya yang memiliki nilai cosinus diantara  $-0.3$  hingga  $-0.5$
- Indikator kemiskinan yang memiliki korelasi rendah atau tidak berkorelasi adalah variabel persentase penduduk yang mengalami gizi buruk dengan variabel persentase penduduk dengan usia sekolah hingga SD/Sederajat dan pasangan lainnya yang memiliki nilai cosinus  $< -0.5$

**Nilai Objek Pada Suatu Objek**

Berdasarkan hasil biplot, dapat diperoleh informasi untuk menentukan indikator kemiskinan di setiap kecamatan. Setiap kecamatan yang terletak searah dengan **vektor** indikator kemiskinan menunjukkan tingginya nilai indikator kemiskinan untuk wilayah tersebut atau dapat diinterpretasikan bahwa indikator kemiskinan untuk wilayah tersebut mempunyai nilai diatas rata-rata seluruh kecamatan. Sebaliknya, jika suatu wilayah terletak berlawanan arah dengan **vektor** indikator kemiskinan maka nilai indikator kemiskinannya rendah atau dibawah rata-rata seluruh kecamatan. Berdasarkan pada Gambar 3. diperoleh bahwa :

- Kelompok 1 terdiri dari kecamatan yang memiliki nilai variabel rasio murid terhadap sekolah pada SMA/SMK/MA dibawah rata-rata.
- Kelompok 2 terdiri dari kecamatan yang memiliki nilai diatas rata-rata untuk setiap variabelnya kecuali pada variabel rasio murid terhadap sekolah pada SMA/SMK/MA.
- Kelompok 3 terdiri dari kecamatan yang memiliki nilai variabel rasio murid terhadap sekolah pada SMA/SMK/MA diatas rata-rata.
- Kelompok 4 terdiri dari kecamatan yang memiliki nilai dibawah rata-rata untuk setiap variabelnya kecuali pada variabel rasio murid terhadap sekolah pada SMA/SMK/MA.

Kemudian dilakukan analisis *K-medoids clustering* dengan matriks G sebanyak 2 dimensi. Pertama ditentukan terlebih dahulu jumlah cluster yang optimum menggunakan metode *Elbow* dan didapatkan grafik berikut dengan menggunakan *Software R*.



**Gambar 4** Jumlah Cluster Optimal

Dari grafik pada **Gambar 4**. dapat dilihat bahwa titik  $k = 3$  membentuk *Elbow* (siku) dimana pada titik tersebut mengalami penurunan nilai *Sum of Square Error* yang signifikan sementara titik selanjutnya memiliki nilai yang konstan, sehingga didapatkan bahwa jumlah cluster yang optimum adalah sebanyak tiga cluster.

Setelah mendapatkan nilai  $k$  atau jumlah cluster yang optimum dengan metode *Elbow* maka akan dilakukan analisis *k-medoids clustering* menggunakan data dari koordinat biplot yang telah dihasilkan pada analisis biplot sebelumnya. Dengan menggunakan *Software R* didapatkan nilai medoid yang digunakan dalam proses *clustering* adalah sebagai berikut :

Tabel 4 Nilai Medoids

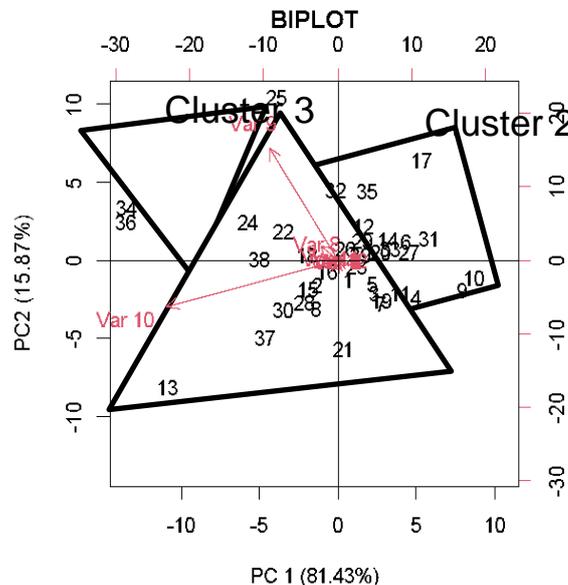
Kecamatan	Nilai Medoid	
	Dim1	Dim2
Puspahiang	-1.935452	-1.7360649
Cineam	2.710724	0.5767774
Rajapolah	-13.459295	3.4634137

Berdasarkan *Software R* didapatkan hasil pengclusteran sebagai berikut :

Tabel 5 Anggota Cluster

Cluster	Anggota Kecamatan	Jumlah Kecamatan
1	Cipatujah, Karangnunggal, Bantarkalong, Taraju, Puspahiang, Tanjungjaya, Salopa, Karangjaya, Manonjaya, Singaparna, Leuwisari, Sariwangi, Kadipaten, dan Pagerageung.	14
2	Cikalong, Pancatengah, Cikatomas, Cibalong, Parungponteng, Bojongasih, Culamega, Bojonggambir, Sodonghilir, Salawu, Sukaraja, Jatiwaras, Cineam, Gunungtanjung, Sukarame, Cigalontang, Padakembang, Sukaratu, Cisayong, Sukahening, Jamanis, dan Sukaresik.	22
3	Mangunreja, Rajapolah, dan Ciawi.	3

Dari **Tabel 5**, dimana terdapat tiga *cluster* yang dapat dilihat berdasarkan indikator kemiskinan dari masing-masing *cluster* dari hasil analisis biplot sebelumnya yang Digambar pada plot berikut :



Gambar 5 Jumlah Cluster dalam Biplot

Masing-masing *cluster* memiliki indikator kemiskinan yang berbeda. Jika dilihat dari **Gambar 5**. Indikator dari setiap *cluster* didapatkan dari nilai setiap objek atau titik koordinat terhadap **vektor** variabel. Sehingga dapat ditentukan objek pada *cluster* tersebut memiliki nilai yang rendah, rata-rata atau tinggi terhadap suatu variabel. Berikut merupakan nilai objek terhadap variabel yang diuraikan menjadi karakteristik dari setiap *cluster* :

- *Cluster 1* berisi kecamatan yang memiliki karakteristik berupa tingginya rasio murid terhadap sekolah pada SMA/SMK/MA.
- *Cluster 2* berisi kecamatan yang memiliki karakteristik yang menengah hingga rendah pada semua variabel kecuali rasio murid terhadap sekolah pada SMP/MTs dan rasio murid terhadap sekolah pada SMA/SMK/MA.
- *Cluster 3* berisi kecamatan yang memiliki karakteristik berupa tingginya rasio murid terhadap sekolah pada SMP/MTs.

**V. KESIMPULAN**

Berdasarkan pembahasan pada bab sebelumnya diperoleh bahwa data berdistribusi normal multivariate. Objek dikelompokkan berdasarkan nilai yang didapat dengan nilai 0 meresentasikan rata-ratanya, sehingga objek dapat dikelompokkan dengan nilai 0 sebagai pemisah antar kelompok. Persentase keragaman data pada biplot sebesar dengan

2 dimensi sehingga dikatakan cukup baik dalam menggambarkan informasi dari data sebenarnya dan dapat memadai dalam pengambilan keputusan.

Dari penelitian ini dihasilkan 3 *cluster* yang memiliki kemiripan dalam indikator kemiskinan. *Cluster 1* terdiri dari 14 kecamatan yaitu Kecamatan Cipatujah, Karangnunggal, Bantarkalong, Taraju, Puspahiang, Tanjungjaya, Salopa, Karangjaya, Manonjaya, Singaparna, Leuwisari, Sariwangi, Kadipaten, dan Pagerageung yang memiliki karakteristik berupa tingginya rasio murid terhadap sekolah pada SMA/SMK/MA. *Cluster 2* terdiri dari 22 kecamatan yaitu Kecamatan Cicalong, Pancatengah, Cikatomas, Cibalong, Parungponteng, Bojongasih, Culamega, Bojonggambir, Sodonghilir, Salawu, Sukaraja, Jatiwaras, Cineam, Gunungtanjung, Sukarame, Cigalontang, Padakembang, Sukaratu, Cisayong, Sukahening, Jamanis, dan Sukaresik yang memiliki karakteristik yang menengah hingga rendah pada semua variabel kecuali variabel rasio murid terhadap sekolah pada SMP/MTs dan rasio murid terhadap sekolah pada SMA/SMK/MA. Sedangkan *Cluster 3* terdiri dari 3 kecamatan yaitu Kecamatan Mangunreja, Rajapolah, dan Ciawi yang memiliki karakteristik berupa tingginya rasio murid terhadap sekolah pada SMP/MTs.

## DAFTAR PUSTAKA

- [1] N. Nurwati, "Kemiskinan : Model Pengukuran, Permasalahan dan Alternatif Kebijakan," Jurnal Kependudukan Padjadjaran, vol. 10, no. 1, 2008.
- [2] Badan Pusat Statistik Jawa Barat, "Jumlah Penduduk Miskin (Ribu Jiwa), 2019-2021." Badan Pusat Statistik Provinsi Jawa Barat. [Online]. Available: <https://jabar.bps.go.id/indicator/23/83/1/jumlah-penduduk-miskin.html>
- [3] UNDP, "The 2020 Global Multidimensional Poverty Index (MPI)," The Oxford Poverty and Human Development Initiative (OPHI), 2020.
- [4] I. T. Jolliffe, Principal component analysis, 2nd ed. in Springer series in statistics. New York: Springer, 2002.
- [5] A. A. Mattjik and I. M. Sumertajaya, "Sidik Peubah Ganda dengan Menggunakan SAS," 2011.
- [6] I. Ginanjar, U. S. Pasaribu, and S. W. Indratno, "A measure for objects clustering in principal component analysis biplot: A case study in inter-city buses maintenance cost data," presented at the STATISTICS AND ITS APPLICATIONS: Proceedings of the 2nd International Conference on Applied Statistics (ICAS II), 2016, Jawa Barat, Indonesia, 2017, p. 020016. doi: 10.1063/1.4979432.
- [7] S. Sharma, Applied multivariate techniques. New York: J. Wiley, 1996.
- [8] T. Purwandari, I. Ginanjar, and D. D. Dewi, "Multiple Correspondence Analysis for Identifying the Contribution of Infant Mortality Indicator Categories," J. Phys.: Conf. Ser., vol. 1776, no. 1, p. 012064, Feb. 2021, doi: 10.1088/1742-6596/1776/1/012064.
- [9] A. C. Rencher and W. F. Christensen, Methods of multivariate analysis, Third Edition. in Wiley series in probability and statistics. Hoboken, New Jersey: Wiley, 2012.
- [10] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, Cluster Analysis, 1st ed. in Wiley Series in Probability and Statistics. Wiley, 2011. doi: 10.1002/9780470977811.
- [11] J. Han and M. Kamber, Data mining: concepts and techniques, 3rd ed. Burlington, MA: Elsevier, 2012.
- [12] J. F. Hair, Ed., Multivariate data analysis, 7th ed. Upper Saddle River, NJ: Prentice Hall, 2010.
- [13] S.-C. Chu, J. Roddick, and J.-S. Pan, "Efficient k-medoids algorithms using multi-centroids with multi-runs sampling scheme," presented at the The Sixth Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2002.
- [14] W. As, M. K. Aidid, and M. Nusrang, "Pengelompokan Kabupaten/Kota Provinsi Sulawesi Selatan dan Barat Berdasarkan Angka Partisipasi Pendidikan SMA/SMK/MA Menggunakan K-Medoid dan CLARA," j. variansi, vol. 1, no. 3, p. 48, Dec. 2019, doi: 10.35580/variansiunm12899.
- [15] A. T. Rahman, "Coal Trade Data Clustering Using K-Means (Case Study Pt. Global Bangkit Utama)," 2017.



© 2023 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (<http://creativecommons.org/licenses/by-sa/4.0/>).