

## Analisis Kluster Metode *Average Linkage* untuk Pengelompokan Kabupaten/Kota Di Provinsi Nusa Tenggara Timur Berdasarkan Tingkat Kesejahteraan Rakyat

Theresia A. Ta'a<sup>1\*</sup>, Elisabeth B. Sinu<sup>2</sup>, Farly O. Haning<sup>3</sup>, Robertus D. Guntur<sup>4</sup>

<sup>1,2,3,4</sup>Program Studi Matematika, Fakultas Sains dan Teknik, Universitas Nusa Cendana

Email: [aidahaidahh76@gmail.com](mailto:aidahaidahh76@gmail.com)

### Article Info

#### Sejarah artikel

Diterima: 20-04-2025

Direvisi: 30-04-2025

Dipublikasikan: 16-05-2025

#### Article history

Received: 20-04-2025

Revised: 30-04-2025

Published: 16-05-2025

### ABSTRAK/ABSTRACT

Kesejahteraan rakyat merupakan salah satu tujuan pembangunan suatu negara. Dalam melaksanakan pembangunan perlu adanya pengidentifikasian karakteristik berdasarkan tingkat kesejahteraan rakyat untuk tiap daerah. Analisis kluster merupakan salah satu teknik multivariat yang dapat digunakan untuk mengelompokkan objek berdasarkan kesamaan karakteristiknya. Penelitian ini bertujuan memberikan gambaran umum tingkat kesejahteraan rakyat di Provinsi Nusa Tenggara Timur dan mengelompokkan 22 kabupaten/kota yang ada di Provinsi Nusa Tenggara Timur berdasarkan tingkat kesejahteraan menggunakan metode *Average Linkage*. Hasil penelitian menunjukkan bahwa tingkat kesejahteraan rakyat di Provinsi Nusa Tenggara Timur tahun 2023 masih belum merata. Lebih lanjut, berdasarkan hasil analisis terlihat bahwa Kota Kupang menduduki peringkat tertinggi di delapan variabel, sedangkan Kabupaten TTS menduduki peringkat tertinggi di dua variabel lainnya. Pengelompokan menggunakan metode *Average Linkage* diperoleh 4 kluster yang terbentuk yakni kluster 1 yang terdiri dari 6 anggota, kluster 2 terdiri dari 13 anggota, kluster 3 terdiri dari 2 anggota, sedangkan kluster 4 terdiri dari 1 anggota.

**Kata Kunci:** Analisis Kluster, *Average Linkage*, Kesejahteraan Rakyat.

#### Abstract

*The welfare of the people is one of the goals of a country's development. In carrying out development, it is necessary to identify the characteristics based on the level of people's welfare for each region. Cluster analysis is one of the multivariate techniques that can be used to group objects based on their characteristic similarities. This study aims to provide a general overview of the level of people's welfare in East Nusa Tenggara Province and to group the 22 regencies/cities in East Nusa Tenggara Province based on their welfare levels using the Average Linkage method. The results of the study indicate that the level of people's welfare in East Nusa Tenggara Province in 2023 is still uneven. Furthermore, the analysis results show that Kupang City ranks highest in eight variables, while TTS Regency ranks highest in the other two variables. Using the Average Linkage method, four clusters were formed: cluster 1 consists of 6 members, cluster 2 consists of 13 members, cluster 3 consists of 2 members, and cluster 4 consists of 1 member.*

**Keywords:** Cluster Analysis, *Average Linkage*, People's Welfare

#### Corresponding Author:

Nama: Theresia A. Ta'a

Email: [aidahaidahh76@gmail.com](mailto:aidahaidahh76@gmail.com)

This is an open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license



## A. Pendahuluan

Kesejahteraan rakyat menjadi tujuan utama dalam pembangunan suatu negara. Seperti halnya di Indonesia, kesejahteraan rakyat adalah salah satu tujuan negara yang tertuang dalam Pembukaan UUD 1945 alinea ke-IV. Dalam melaksanakan program pembangunan perlu adanya pengidentifikasian karakteristik berdasarkan tingkat kesejahteraan rakyat setiap daerah agar pemerintah dalam mengambil kebijakan dan strategi pembangunan dapat sesuai dengan tujuan negara itu sendiri. Menurut Badan Pusat Statistik terdapat delapan indikator yang dapat digunakan untuk mengukur tingkat kesejahteraan rakyat, yaitu indikator kependudukan, kesehatan dan gizi, taraf dan pola konsumsi, ketenagakerjaan, pendidikan, perumahan dan lingkungan, kemiskinan dan sosial lainnya [1].

Jumlah penduduk Provinsi Nusa Tenggara Timur tahun 2023 mencapai 5,57 juta jiwa dengan tingkat kepadatan penduduk sebesar 120 jiwa/km<sup>2</sup> [2]. Jumlah penduduk yang padat dan distribusi yang tidak merata dapat menimbulkan masalah seperti tekanan penduduk yang nantinya akan mempengaruhi daya dukung pembangunan infrastruktur, lingkungan serta sosial di wilayah tertentu. Selain masalah kependudukan, indikator lain yang mempengaruhi tingkat kesejahteraan rakyat adalah kemiskinan. Provinsi Nusa Tenggara Timur menduduki peringkat ketiga dengan persentase penduduk miskin paling banyak di Indonesia sebesar 19,96% [1]. Sesuai data tahun 2023, jumlah penduduk miskin meningkat mencapai 1,14 juta jiwa [3]. Jumlah penduduk miskin akan semakin meningkat jika tidak ada perubahan strategi dalam mengatasi kemiskinan.

Berdasarkan permasalahan tersebut maka perlu dilakukan pengidentifikasian karakteristik berdasarkan tingkat kesejahteraan tiap daerah sehingga pemerintah dapat mengambil kebijakan dan strategi yang baik dalam melaksanakan pembangunan yang merata dan tepat sasaran. Analisis kluster merupakan salah satu teknik multivariat yang dapat digunakan untuk mengelompokkan objek berdasarkan kesamaan karakteristiknya [4]. Objek yang memiliki tingkat kemiripan tinggi akan berada dalam satu kluster, sedangkan objek dengan tingkat kemiripan rendah akan berbeda kluster. Adapun penelitian terdahulu yang menggunakan analisis kluster dalam penelitiannya seperti Cesario et al (2022) [5], Latuhimallo et al. (2023) [6] dan Damayanthi et al. (2023) [7].

Berdasarkan penelitian yang dilakukan oleh Reinaldi et al. (2021) mengenai Perbandingan Metode Single Linkage, Complete Linkage, dan Average Linkage Pada Kesejahteraan Masyarakat di Jawa Timur menunjukkan bahwa metode Average Linkage merupakan metode terbaik dari ketiga metode yang diteliti [8]. Adapun penelitian lainnya yang dilakukan Ramadani & Salma (2022) mengenai perbandingan Metode Average Linkage dan Ward dalam Pengelompokan Kesejahteraan Sumatera Barat tahun 2021 menunjukkan bahwa metode Average Linkage merupakan metode terbaik dibandingkan metode Ward [9].

Penelitian ini bertujuan untuk mendeskripsikan gambaran umum tingkat kesejahteraan rakyat kabupaten/kota di Provinsi Nusa Tenggara Timur tahun 2023 serta mengelompokkan kabupaten/kota di Provinsi Nusa Tenggara Timur berdasarkan tingkat kesejahteraan rakyat dengan analisis kluster menggunakan metode *Average Linkage*. Objek penelitian ini adalah 22 kabupaten/kota di Provinsi Nusa Tenggara Timur. Variabel yang digunakan yaitu kepadatan penduduk (X<sub>1</sub>), umur harapan hidup (X<sub>2</sub>), pengeluaran riil perkapita (X<sub>3</sub>), PDRB atas dasar harga berlaku (X<sub>4</sub>), jumlah angkatan kerja (X<sub>5</sub>), angka melek huruf (X<sub>6</sub>), rata-rata lama sekolah (X<sub>7</sub>), rumah tangga dengan sumber penerangan PLN (X<sub>8</sub>), jumlah penduduk miskin (X<sub>9</sub>), dan tingkat pengangguran terbuka (X<sub>10</sub>).

## B. Metode Penelitian

Jenis data yang digunakan pada penelitian ini adalah data sekunder yaitu data yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur berupa data kepadatan penduduk (X<sub>1</sub>), umur harapan hidup (X<sub>2</sub>), pengeluaran riil perkapita (X<sub>3</sub>), PDRB atas dasar harga berlaku (X<sub>4</sub>), jumlah angkatan kerja (X<sub>5</sub>), angka melek huruf (X<sub>6</sub>), rata-rata lama sekolah (X<sub>7</sub>), rumah tangga dengan sumber penerangan PLN (X<sub>8</sub>), jumlah penduduk miskin (X<sub>9</sub>), dan tingkat pengangguran terbuka (X<sub>10</sub>). Objek penelitian mencakup 22 kabupaten/kota yang ada di Provinsi Nusa Tenggara Timur.

Prosedur penelitian ini adalah sebagai berikut:

1. Melakukan pengumpulan data berdasarkan variabel yang digunakan pada penelitian ini.
2. Melakukan analisis statistik deskriptif
3. Melakukan standarisasi data menggunakan persamaan (1) [10]:

$$z = \frac{x - \bar{x}}{s} \quad (1)$$

keterangan:

x = nilai data

$\bar{x}$  = nilai rata-rata

s = standar deviasi

4. Melakukan pengujian asumsi. Terdapat dua asumsi pada analisis kluster, yaitu:

a. Uji Kecukupan Sampel, yaitu dengan menggunakan *Kaiser Meyer Olkin* (KMO). Nilai KMO dapat dihitung menggunakan persamaan (2) [11]:

$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}}{\sum_{i=1}^p \sum_{j=1}^p r_{ij} + \sum_{i=1}^p \sum_{j=1}^p a_{ij}} \quad (2)$$

keterangan:

$r_{ij}^2$  = koefisien korelasi sederhana antara variabel i dan j

$a_{ij}^2$  = koefisien korelasi parsial antara variabel i dan j

- b. Uji Multikolinearitas, yaitu dengan melihat nilai korelasi *pearson* [12]. Nilai korelasi *pearson* dapat dihitung menggunakan persamaan (3) :

$$r_{x_i x_j} = \frac{n \sum x_i x_j - \sum x_i \sum x_j}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum x_j^2 - (\sum x_j)^2}} \quad (3)$$

keterangan:

$r_{x_i x_j}$  = koefisien korelasi antara variabel

$n$  = banyaknya pasangan data

$\sum x_i$  = total jumlah dari variabel i

$\sum x_j$  = total jumlah dari variabel j

$\sum x_i x_j$  = hasil perkalian dari total jumlah variabel i dan j

5. Menentukan ukuran kemiripan objek dengan jarak *Euclidian* [13].

Pengukuran jarak *Euclidian* dapat menggunakan persamaan (4):

$$d_{(i,j)} = \sqrt{(x_{i1} - x_{j1})^2 + \dots + (x_{ip} - x_{jp})^2} \quad (4)$$

keterangan:

$d_{(i,j)}$  = jarak antara objek ke-i dan objek ke-j

$x_{ip}$  = nilai objek ke-i pada peubah ke-p

$x_{jp}$  = nilai objek ke-j pada peubah ke-p

$p$  = banyaknya peubah yang diamati

6. Melakukan proses analisis kluster dengan metode *Average Linkage*. *Average Linkage* adalah metode pengelompokan yang melihat jarak rata-rata antar objek (*average distance*). Adapun tahapan dalam melakukan pengklasteran, yaitu:

a. Menentukan objek bersesuaian yang memiliki jarak terdekat dalam matriks  $D = \{d_{ik}\}$ ,

b. Menggabungkan objek yang bersesuaian, misalkan objek  $U$  dan objek  $V$ , yang kemudian didapatkan kluster ( $UV$ ),

c. Menghitung jarak antara kluster ( $UV$ ) dengan objek lain misalkan  $W$  yang belum bergabung menggunakan persamaan (5):

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{(UV)} N_W} \quad (5)$$

Keterangan

$d_{ik}$  = jarak antara objek i pada kluster  $UV$  dan objek k pada kluster  $W$

$N_{(UV)}$  = jumlah objek pada kluster  $UV$

$N_W$  = jumlah objek pada kluster  $UV$  dan  $W$

dan menghitung kembali matriks jarak baru dengan cara pada langkah ketiga.

d. Ulangi langkah dua sampai semua objek bergabung menjadi satu kluster [14].

7. Menentukan jumlah kluster dan anggotanya.

8. Interpretasi Kluster. Interpretasi Dilakukan dengan membandingkan rata-rata (*centroid*) pada setiap kelompok. Nilai centroid dapat dihitung menggunakan persamaan (6) [12]:

$$C = \frac{\sum_{j=1}^n \sum_{k=1}^p x_{jk}}{N} \quad (6)$$

keterangan:

$C$  = nilai rata-rata (*centroid*) kluster

$x_{jk}$  = nilai atau data dari objek j pada variabel k dalam kluster yang dicari nilai centroid

N = Jumlah objek dalam kluster yang dicari nilai centroid

## C. Hasil dan Pembahasan

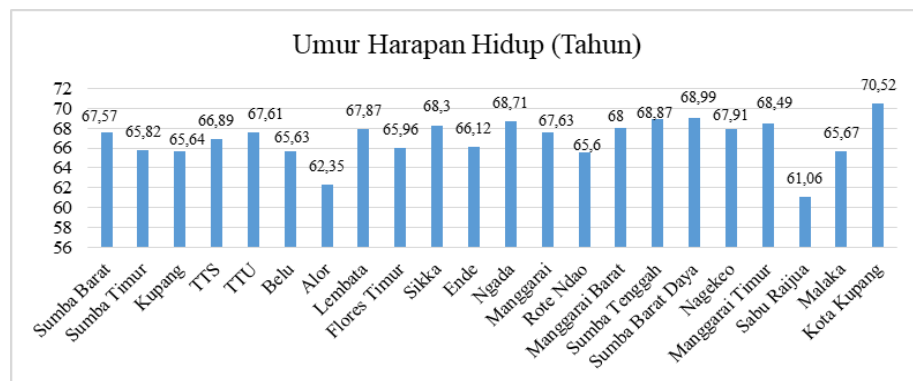
### 1.1 Statistik Deskriptif

Statistik deskriptif digunakan untuk mengetahui karakteristik data untuk masing-masing variabel yang digunakan. Berikut adalah grafik yang menggambarkan kondisi tingkat kesejahteraan rakyat di Provinsi Nusa Tenggara Timur tahun 2023 yang ditampilkan pada Gambar 1 sampai Gambar 10.



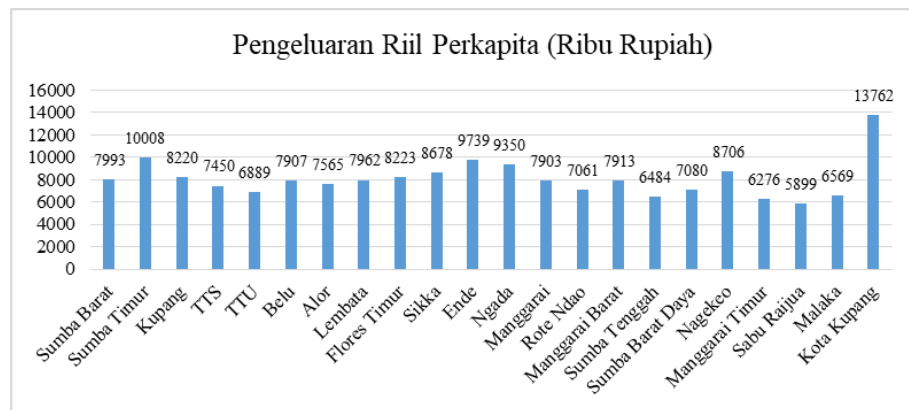
**Gambar 1.** Grafik Kepadatan Penduduk

Berdasarkan Gambar 2 dapat dilihat bahwa Kota Kupang merupakan daerah dengan tingkat kepadatan penduduk tertinggi yaitu sebanyak 2.929 jiwa/km<sup>2</sup>, sedangkan terendah adalah Kabupaten Sumba Timur sebanyak 37 jiwa/km<sup>2</sup>.



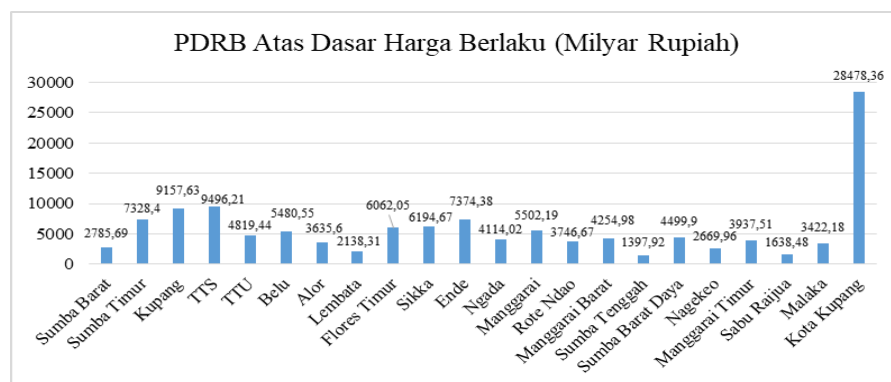
**Gambar 2.** Grafik Umur Harapan Hidup

Berdasarkan Gambar 2 menunjukkan bahwa Kota Kupang merupakan daerah dengan umur harapan hidup tertinggi yaitu sebanyak 70,52 tahun, sedangkan terendah ditempati oleh Kabupaten Sabu Raijua yaitu sebanyak 61,06 tahun.



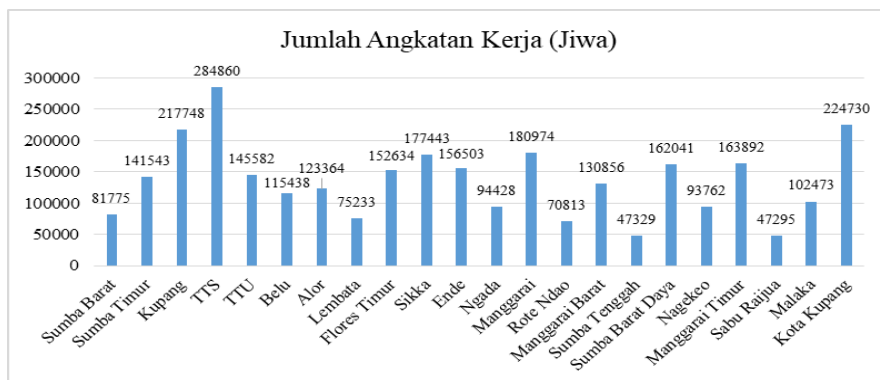
**Gambar 3.** Grafik Pengeluaran Riil Perkapita

Berdasarkan Gambar 3 terlihat bahwa Kota Kupang merupakan daerah dengan nilai pengeluaran riil perkapita tertinggi yaitu sebesar 13,762 juta rupiah, sedangkan terendah adalah Kabupaten Sabu Raijua sebesar 5,899 juta rupiah.



**Gambar 4.** Grafik PDRB Atas Dasar Harga Berlaku

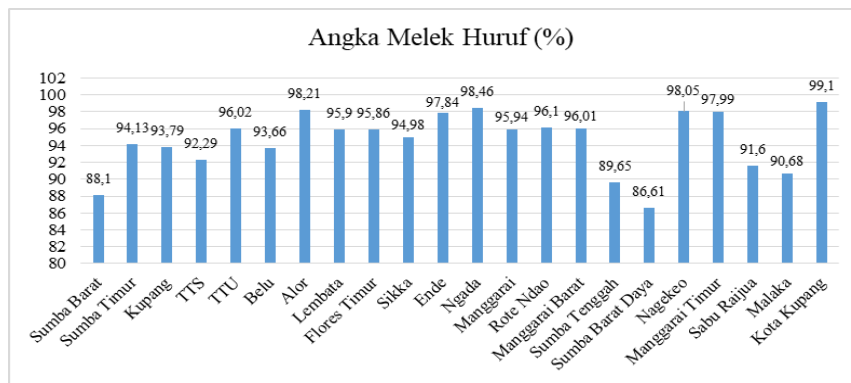
Berdasarkan Gambar 4 terlihat bahwa Kota Kupang merupakan daerah dengan nilai PDRB atas dasar harga berlaku tertinggi yaitu sebesar 28,478 triliun rupiah, sedangkan terendah Kabupaten Sumba Tengah sebesar 1,397 triliun rupiah.



**Gambar 5.** Grafik Jumlah Angkatan Kerja

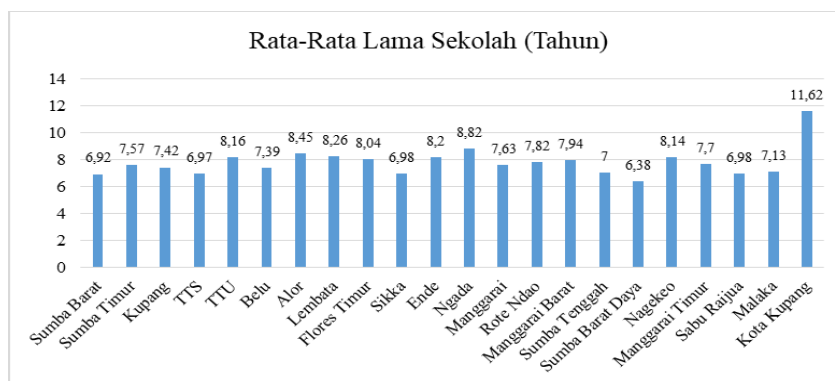
Berdasarkan Gambar 5 terlihat bahwa Kabupaten Timor Tengah Selatan (TTS) memiliki jumlah angkatan kerja tertinggi yaitu sebanyak 284.860 jiwa, sedangkan terendah adalah Kabupaten Sabu Raijua sebanyak 47.295 jiwa.





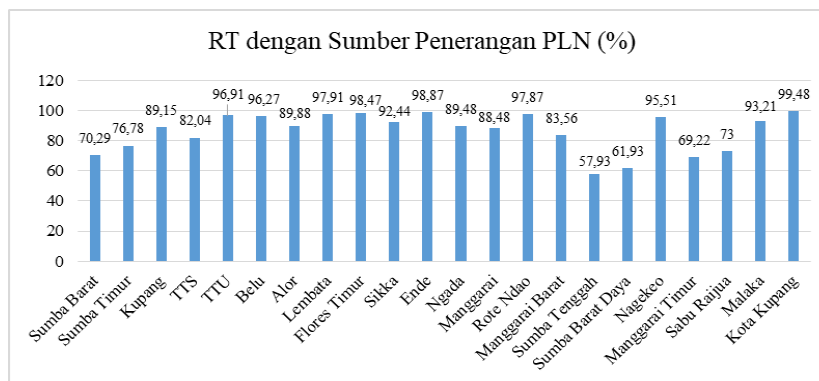
**Gambar 6.** Grafik Angka Melek Huruf

Berdasarkan Gambar 6 terlihat bahwa Kota Kupang merupakan daerah dengan angka melek huruf tertinggi yaitu sebesar 99,1%, sedangkan terendah adalah Kabupaten Sumba Barat Daya yaitu sebesar 86,61%.



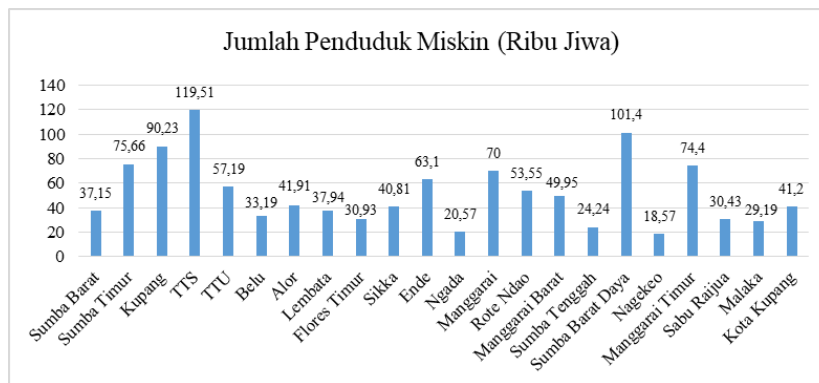
**Gambar 7.** Grafik Rata-Rata Lama Sekolah

Berdasarkan Gambar 7 terlihat bahwa Kota Kupang merupakan daerah dengan rata-rata lama sekolah tertinggi yaitu sebesar 11,61 tahun, sedangkan terendah adalah Kabupaten Sumba Barat Daya yaitu sebesar 6,38 tahun.



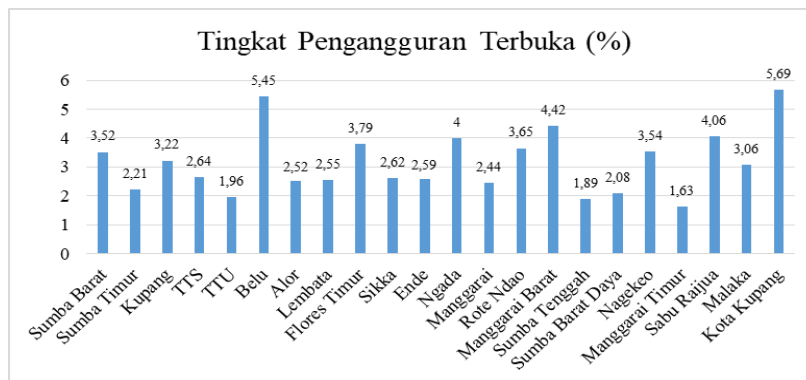
**Gambar 8.** Grafik Persentase RT dengan Sumber Penerangan PLN

Berdasarkan Gambar 8 terlihat bahwa Kota Kupang merupakan daerah dengan jumlah rumah tangga dengan sumber penerangan PLN tertinggi yaitu sebanyak 99,48%, sedangkan terendah adalah Kabupaten Sumba Tengah yaitu sebesar 57,93%.



**Gambar 9.** Grafik Jumlah Penduduk Miskin

Berdasarkan Gambar 9 terlihat bahwa Kabupaten Timor Tengah Selatan (TTS) merupakan daerah dengan jumlah penduduk miskin tertinggi yaitu sebanyak 119,51 ribu jiwa, sedangkan terendah adalah Kabupaten Nagekeo yaitu sebanyak 18,57 ribu jiwa.



**Gambar 10.** Grafik Tingkat Pengangguran Terbuka

Berdasarkan Gambar 10 terlihat bahwa Kota Kupang merupakan daerah dengan tingkat pengangguran terbuka tertinggi yaitu sebanyak 5,69%, sedangkan terendah adalah Kabupaten Manggarai Timur yaitu sebesar 1,63%.

## 1.2 Standarisasi Data

Proses standarisasi data digunakan jika terdapat perbedaan satuan yang signifikan antara variabel. Karena dalam penelitian ini data yang digunakan memiliki skala satuan yang berbeda maka standarisasi data terlebih dahulu. Hasil standarisasi data ditampilkan pada Tabel 1.

**Tabel 1.** Standarisasi Data

| No | Kabupaten/Kota | ZX <sub>1</sub> | ZX <sub>2</sub> | ZX <sub>3</sub> | ... | ZX <sub>10</sub> |
|----|----------------|-----------------|-----------------|-----------------|-----|------------------|
| 1  | Sumba Barat    | -<br>0,10494    | 0,32318         | -<br>0,04889    | ... | 0,32742          |
| 2  | Sumba Timur    | -<br>0,37911    | -0,48845        | 1,16121         | ... | -0,86554         |



|    |             |              |          |              |     |          |
|----|-------------|--------------|----------|--------------|-----|----------|
| 3  | Kupang      | -<br>0,31893 | -0,57194 | 0,08743      | ... | 0,05423  |
| 4  | TTS         | -<br>0,23868 | 0,00780  | -<br>0,37499 | ... | -0,47396 |
| 5  | TTU         | -<br>0,26878 | 0,34173  | -<br>0,71189 | ... | -1,09321 |
| ⋮  | ⋮           | ⋮            | ⋮        | ⋮            | ⋮   | -0,58324 |
| 22 | Kota Kupang | 4,45567      | 1,69136  | 3,41566      | ... | 2,30356  |

Berdasarkan Tabel 1 dapat ditunjukkan bahwa ZX1, ZX2, sampai ZX10 adalah variabel yang datanya telah distandarisasi untuk setiap kabupaten/kota.

### 1.3 Uji Asumsi Kluster

#### a. Uji kecukupan Sampel

Untuk mengetahui apakah sampel yang digunakan telah cukup untuk dianalisis, dapat dilihat dari nilai *Kaiser Meyer Olkin* (KMO) yang ditampilkan pada Tabel 2.

**Tabel 2.** Nilai KMO dan Bartlett Test

| KMO and Bartlett's Test       |                   |            |         |    |       |
|-------------------------------|-------------------|------------|---------|----|-------|
| Kaiser-Meyer-Olkin            |                   |            | 0,656   |    |       |
| Bartlett's Test of Sphericity | Approx. Square df | Chi-Square | 185,282 | 45 | 0,000 |
|                               |                   |            |         |    |       |

Tabel 2 menunjukkan bahwa nilai KMO lebih besar dari 0,5 yaitu sebesar 0,656 artinya sampel telah cukup dan layak untuk dilakukan analisis kluster.

#### b. Uji Multikolinearitas

Untuk mengetahui terjadinya multikolinearitas dapat dilihat pada matriks korelasi di Tabel 3. Dikatakan multikolinearitas jika nilai korelasinya lebih besar dari 0,80.

**Tabel 3.** Korelasi antar variabel

| Variabel       | X <sub>1</sub> | X <sub>2</sub> | X <sub>3</sub> | X <sub>4</sub> | X <sub>5</sub> | ... | X <sub>10</sub> |
|----------------|----------------|----------------|----------------|----------------|----------------|-----|-----------------|
| X <sub>1</sub> | 1              | 0,375          | 0,746          | 0,909          | 0,336          | ... | 0,530           |
| X <sub>2</sub> | 0,375          | 1              | 0,394          | 0,347          | 0,263          | ... | -0,044          |

|          |        |        |        |       |        |     |        |
|----------|--------|--------|--------|-------|--------|-----|--------|
| $X_3$    | 0,746  | 0,394  | 1      | 0,822 | 0,381  | ... | 0,466  |
| $X_4$    | 0,909  | 0,347  | 0,822  | 1     | 0,636  | ... | 0,450  |
| $X_5$    | 0,336  | 0,263  | 0,381  | 0,636 | 1      | ... | -0,048 |
| $X_6$    | 0,253  | 0,061  | 0,460  | 0,312 | 0,165  | ... | 0,198  |
| $X_7$    | 0,784  | 0,288  | 0,776  | 0,735 | 0,186  | ... | 0,480  |
| $X_8$    | 0,229  | -0,079 | 0,419  | 0,312 | 0,165  | ... | 0,407  |
| $X_9$    | -0,088 | 0,072  | -0,072 | 0,194 | 0,722  | ... | -0,415 |
| $X_{10}$ | 0,530  | -0,044 | 0,466  | 0,450 | -0,048 | ... | 1      |

Tabel 3 menunjukkan bahwa variabel  $X_1$  dengan  $X_4$  dan variabel  $X_3$  dengan  $X_4$  mengalami multikolinearitas karena nilai korelasi lebih dari 0,80. Karena terjadi multikolinearitas, maka perlu dilakukan penanganan guna memenuhi asumsi dalam analisis kluster. Penanganan masalah multikolinearitas dilakukan dengan *Principal Component Analysis (PCA)* [15].

#### 1.4 *Principal Component Analisis (PCA)*

Dalam *PCA* terdapat beberapa asumsi yang harus terpenuhi yaitu:

##### a. *Uji Kecukupan Sampel*

Uji kecukupan sampel menggunakan *KMO* telah dilakukan sebelumnya dan diperoleh nilai *KMO* sebesar 0,656 dan telah memenuhi asumsi.

##### b. *Uji Bartlett*

Uji Bartlett bertujuan untuk mengetahui apakah terdapat korelasi antar variabel. Hasil pengujian dapat dilihat pada Tabel 3, terlihat bahwa nilai Bartlett menunjukkan nilai *p-value* = 0,000 lebih kecil dari 0,05 sehingga dapat disimpulkan bahwa terdapat korelasi antar variabel.

##### c. *Nilai Measure of Sampling Adequacy (MSA)*

Nilai *MSA* digunakan untuk mengetahui variabel-variabel mana saja yang layak untuk dilakukan *PCA*. Nilai *MSA* yang lebih dari 0,5 menunjukkan bahwa variabel tersebut layak untuk dilakukan *PCA*. Hasil perhitungan nilai *MSA* ditampilkan pada Tabel 4 berikut.

| Tabel 4. Nilai <i>MSA</i> |                  |
|---------------------------|------------------|
| Variabel                  | Nilai <i>MSA</i> |
| $X_1$                     | 0,649            |

|                 |       |
|-----------------|-------|
| X <sub>2</sub>  | 0,464 |
| X <sub>3</sub>  | 0,729 |
| X <sub>4</sub>  | 0,614 |
| X <sub>5</sub>  | 0,492 |
| X <sub>6</sub>  | 0,584 |
| X <sub>7</sub>  | 0,738 |
| X <sub>8</sub>  | 0,808 |
| X <sub>9</sub>  | 0,612 |
| X <sub>10</sub> | 0,853 |

Berdasarkan Tabel 4 dapat dilihat nilai MSA dari setiap variabel. Dari sepuluh variabel yang ada, variabel X<sub>2</sub> dan X<sub>5</sub> masing-masing memiliki nilai MSA kurang dari 0,5 sehingga hanya delapan variabel yang layak untuk dilakukan PCA.

Selanjutnya dilakukan PCA. Langkah pertama adalah membuat Matriks Z yang memuat nilai dari variabel yang telah distandarisasi.

$$Z = \begin{bmatrix} -0,10494 & -0,04889 & -0,55070 & -1,86013 & \cdots & 0,32742 \\ -0,37911 & 1,16121 & 0,27259 & -1,13172 & \cdots & -0,86554 \\ -0,31893 & 0,08743 & 0,60411 & -0,22918 & \cdots & 0,05423 \\ -0,23868 & -0,37499 & 0,66547 & -0,65913 & \cdots & -0,47396 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 4,4557 & 3,41566 & 4,10568 & 1,29286 & \cdots & 2,30356 \end{bmatrix}$$

Setelah diperoleh matriks Z selanjutnya membuat matriks korelasi R.

$$R = \begin{bmatrix} 1 & 0,746 & 0,909 & 0,253 & 0,784 & 0,229 & -0,088 & 0,530 \\ 0,746 & 1 & 0,822 & 0,460 & 0,776 & 0,419 & -0,072 & 0,466 \\ 0,909 & 0,822 & 1 & 0,312 & 0,735 & 0,312 & 0,194 & 0,450 \\ 0,253 & 0,460 & 0,312 & 1 & 0,723 & 0,662 & -0,209 & 0,198 \\ 0,784 & 0,776 & 0,735 & 0,723 & 1 & 0,537 & -0,273 & 0,480 \\ 0,229 & 0,419 & 0,312 & 0,662 & 0,537 & 1 & -0,254 & 0,407 \\ -0,088 & -0,072 & 0,194 & -0,209 & -0,273 & -0,254 & 1 & -0,415 \\ 0,530 & 0,194 & 0,450 & 0,198 & 0,480 & 0,407 & -0,415 & 1 \end{bmatrix}$$

Langkah selanjutnya adalah mencari nilai eigen dari matriks korelasi R. Nilai eigen ini dapat menjelaskan besarnya kontribusi keragaman atau varian masing-masing komponen utama.

**Tabel 5.** Nilai Eigen, Total Varian dan Total Kumulatif Varian

| Komponen | Nilai Eigen | Total Varian (%) | Total Kumulatif Varian (%) |
|----------|-------------|------------------|----------------------------|
| 1        | 4,342       | 54,281           | 54,281                     |
| 2        | 1,490       | 18,619           | 72,900                     |
| 3        | 1,055       | 13,191           | 86,091                     |
| 4        | 0,547       | 6,841            | 92,932                     |
| 5        | 0,262       | 3,260            | 96,211                     |
| 6        | 0,229       | 2,867            | 99,078                     |

|   |       |       |         |
|---|-------|-------|---------|
| 7 | 0,054 | 0,671 | 99,750  |
| 8 | 0,020 | 0,250 | 100,000 |

Pada Tabel 5 terlihat bahwa komponen 1, komponen 2, komponen 3 secara bersama-sama telah menjelaskan keragaman dari kedelapan variabel sebesar 86,091 %.

Selanjutnya ditentukan eigen vektor atau koefisien dari principal component yang akan digunakan untuk membentuk persamaan PC.

**Tabel 6.** Koefisien *Principal Component (PC)*

| Variabel        | <i>Principal Component (PC)</i> |                 |                 |
|-----------------|---------------------------------|-----------------|-----------------|
|                 | PC <sub>1</sub>                 | PC <sub>2</sub> | PC <sub>3</sub> |
| X <sub>1</sub>  | 0,845                           | 0,385           | -0,242          |
| X <sub>3</sub>  | 0,875                           | 0,219           | 0,030           |
| X <sub>4</sub>  | 0,843                           | 0,511           | 0,014           |
| X <sub>6</sub>  | 0,642                           | -0,434          | 0,545           |
| X <sub>7</sub>  | 0,938                           | -0,062          | 0,102           |
| X <sub>8</sub>  | 0,621                           | -0,503          | 0,330           |
| X <sub>9</sub>  | -0,237                          | 0,742           | 0,516           |
| X <sub>10</sub> | 0,647                           | -0,189          | -0,559          |

Berdasarkan Tabel 6 di atas dapat dibentuk tiga persamaan sebagai berikut:

$$PC_1 = 0,845ZX_1 + 0,875ZX_3 + 0,843ZX_4 + 0,642ZX_6 + 0,938ZX_7 + 0,621ZX_8 - 0,237ZX_9 + 0,647ZX_{10} \quad (1)$$

$$PC_2 = 0,385ZX_1 + 0,219ZX_3 + 0,511ZX_4 - 0,434ZX_6 - 0,062ZX_7 - 0,503ZX_8 + 0,742ZX_9 - 0,189ZX_{10} \quad (2)$$

$$PC_3 = -0,242ZX_1 + 0,030ZX_3 + 0,014ZX_4 + 0,545ZX_6 + 0,102ZX_7 + 0,330ZX_8 + 0,516ZX_9 - 0,559ZX_{10} \quad (3)$$

Persamaan tersebut akan digunakan untuk mencari nilai atau score PC untuk setiap kabupaten/kota. Hasil perhitungan  $PC_1$ ,  $PC_2$  dan  $PC_3$  untuk semua kabupaten/kota ditampilkan dalam Tabel 7

**Tabel 7.** Koefisien *Principal Component (PC)*

| No | Kabupaten/Kota | PC <sub>1</sub> | PC <sub>2</sub> | PC <sub>3</sub> |
|----|----------------|-----------------|-----------------|-----------------|
| 1  | Sumba Barat    | -3,01526        | 0,69567         | -1,96119        |
| 2  | Sumba Timur    | -0,59414        | 1,50889         | 0,72711         |
| 3  | Kupang         | -0,32907        | 1,25429         | 0,70125         |
| 4  | TTS            | -2,23744        | 2,61095         | 1,06011         |
| 5  | TTU            | -0,65131        | -0,61941        | 1,28763         |
| ⋮  | ⋮              | ⋮               | ⋮               | ⋮               |
| 22 | Kota Kupang    | 16,69144        | 2,52736         | -0,99203        |

## 1.5 Menentukan Ukuran Kemiripan Objek

Semakin besar jarak menunjukkan ketidakmiripan antara objek, begitupun sebaliknya. Ukuran jarak yang digunakan adalah jarak Euclidian. Dihitung jarak antara kabupaten Sumba Barat dan Sumba Timur (objek 1 dan 2):

$$\begin{aligned}
 d_{1,2} &= \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2} \\
 &= \sqrt{(-3,01526 - (-0,59414))^2 + (0,69567 - 1,50889)^2 + (-1,196119 - 0,72711)^2} \\
 &= 3,708
 \end{aligned}$$

Berikut ini merupakan ringkasan matriks dari hasil perhitungan jarak Euclidian untuk seluruh objek

$$d_{i,j} = \begin{bmatrix} 0,000 & 3,708 & 3,823 & 3,661 & 4,228 & \dots & 19,815 \\ 3,708 & 0,000 & 3,368 & 2,006 & 2,202 & \dots & 17,401 \\ 3,823 & 3,368 & 0,000 & 2,369 & 1,990 & \dots & 17,152 \\ 3,661 & 2,006 & 2,369 & 0,000 & 3,606 & \dots & 19,040 \\ 4,228 & 2,202 & 1,990 & 3,606 & 0,000 & \dots & 17,773 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 19,815 & 17,401 & 17,152 & 19,040 & 17,773 & \dots & 0,000 \end{bmatrix}$$

#### 1.6 Proses Kluster dengan Metode *Average Linkage*

Setelah menghitung jarak masing-masing objek terhadap objek lain, selanjutnya adalah melakukan pengklasteran dengan metode *Average Linkage*. Adapun langkah-langkahnya adalah sebagai berikut:

Langkah 1. Menentukan pasangan objek terdekat yaitu Kabupaten Alor (7) dan Kabupaten Lembata (8) dengan jarak sebesar 0,358.

Langkah 2. Menggabungkan pasangan objek tersebut menjadi satu kluster (7,8)

Langkah 3. Menghitung jarak antara kluster (7,8) dengan objek lain yang belum bergabung.

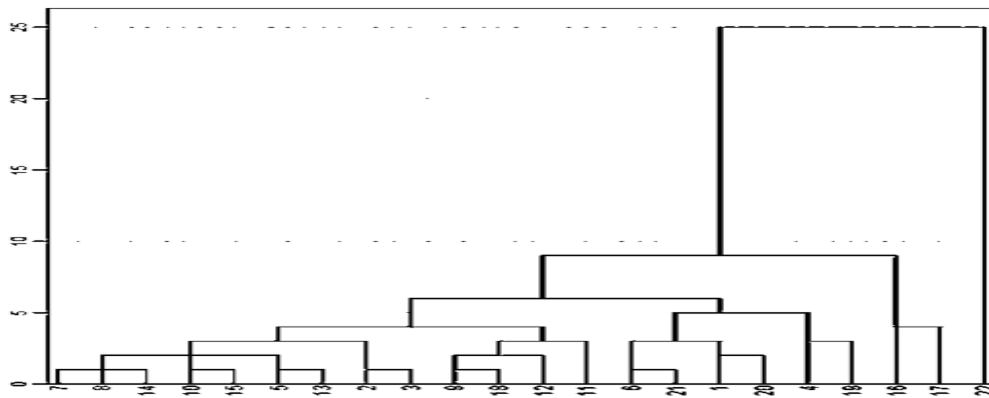
Contoh perhitungan

$$d_{(7,8)1} = \frac{d_{7,1} + d_{8,1}}{2} = \frac{4,755 + 4,594}{2} = 4,675$$

Seterusnya hingga nantinya terbentuk matriks jarak yang baru

Langkah 4. Ulangi setiap langkah sampai akhirnya semua objek bergabung menjadi satu kluster.

Hasil perhitungan ini ditampilkkan dalam Gambar 11 berikut



**Gambar 11.** Dendrogram Metode Average Linkage

### 1.7 Jumlah Kluster dan Anggotanya

Dengan melihat Gambar 11 maka ditentukan banyaknya kluster adalah 4 kluster. Banyaknya anggota masing-masing kluster ditampilkan dalam Tabel 8 berikut:

**Tabel 8.** Kluster Beserta Anggotanya

| Kluster | Kabupaten/Kota                                                                                                            |
|---------|---------------------------------------------------------------------------------------------------------------------------|
| 1       | Sumba Barat, TTS, Belu, Manggarai Timur, Sabu Raijua, Malaka                                                              |
| 2       | Sumba Timur, Kupang, TTU, Alor, Lembata, Flores Timur, Sikka, Ende, Ngada, Manggarai, Rote Ndao, Manggarai Barat, Nagekeo |
| 3       | Sumbah Tengah, Sumba Barat Daya                                                                                           |
| 4       | Kota Kupang                                                                                                               |

### 1.8 Interpretasi kluster

Untuk mempermudah mengetahui karakteristik setiap kluster maka perlu dihitung nilai centroid atau nilai rata-rata tiap kluster yang ditampilkan pada Tabel 9 berikut:

**Tabel 9.** Centroid Kluster

| Variabel       | Klaster 1 | Klaster 2 | Klaster 3 | Klaster 4 |
|----------------|-----------|-----------|-----------|-----------|
| X <sub>1</sub> | -0,15593  | -0,23946  | -0,20358  | 4,45567   |
| X <sub>3</sub> | -0,63583  | 0,150122  | -0,77615  | 3,41566   |
| X <sub>4</sub> | -0,24724  | -0,12154  | -0,52112  | 4,10568   |
| X <sub>6</sub> | -0,63142  | 0,476828  | -1,85154  | 1,29286   |
| X <sub>7</sub> | -0,58651  | 0,152462  | -1,05563  | 3,64829   |
| X <sub>8</sub> | -0,44220  | 0,443085  | -2,07074  | 1,03455   |

|          |         |          |          |          |
|----------|---------|----------|----------|----------|
| $X_9$    | 0,07783 | -0,06781 | 0,40410  | -0,39369 |
| $X_{10}$ | 0,21207 | -0,11039 | -1,07044 | 2,30356  |

#### D. Kesimpulan

Berdasarkan hasil analisis dapat disimpulkan bahwa gambaran umum tingkat kesejahteraan rakyat di provinsi Nusa Tenggara Timur tergolong masih belum merata. Hal ini dapat dilihat dari banyaknya jumlah penduduk miskin di Kabupaten Timor Tengah Selatan (TTS) dan kepadatan penduduk yang tinggi di Kota Kupang. Selain itu, Kota Kupang juga menduduki peringkat tertinggi pada variabel umur harapan hidup, pengeluaran riil perkapita, PDRB atas dasar harga berlaku, angka melek huruf, rata-rata lama sekolah, rumah tangga dengan sumber penerangan PLN dan tingkat pengangguran terbuka. Kemudian, jumlah angkatan kerja yang paling tinggi dimiliki oleh Kabupaten Timor Tengah Selatan. Pengelompokan 22 kabupaten/kota yang ada di provinsi Nusa Tenggara Timur dengan menggunakan analisis klaster metode *average linkage* diperoleh 4 klaster yang terbentuk. Klaster 1 terdiri dari 6 anggota dengan variabel yang mempengaruhi yaitu variabel kepadatan penduduk, jumlah penduduk miskin dan tingkat pengangguran terbuka. Sedangkan untuk kelima variabel lainnya memiliki nilai rata-rata lebih tinggi dibandingkan klaster 3 tetapi lebih rendah dibandingkan dua klaster lainnya. Klaster 2 terdiri dari 13 anggota dengan variabel yang mempengaruhi yaitu variabel angka melek huruf ( $X_6$ ), rumah tangga dengan sumber penerangan PLN ( $X_8$ ), pengeluaran riil perkapita ( $X_3$ ), PDRB atas harga berlaku ( $X_4$ ), dan rata-rata lama sekolah ( $X_7$ ). Selain variabel jumlah penduduk miskin ( $X_9$ ) dan tingkat pengangguran terbuka ( $X_{10}$ ) untuk variabel kepadatan penduduk ( $X_1$ ) memiliki nilai rata-rata yang paling rendah dibandingkan klaster lainnya. Klaster 3 terdiri dari 2 anggota dengan variabel yang mempengaruhi yaitu variabel jumlah penduduk miskin ( $X_9$ ). Selain dari variabel kepadatan penduduk ( $X_1$ ) untuk keenam variabel lainnya memiliki nilai rata-rata yang paling rendah dibandingkan dengan klaster lainnya. Klaster 3 memiliki tingkat kesejahteraan rakyat yang sangat rendah. Klaster 4 terdiri dari 1 anggota dengan variabel yang mempengaruhi yaitu variabel kepadatan penduduk ( $X_1$ ), pengeluaran riil perkapita ( $X_3$ ), PDRB atas harga berlaku ( $X_4$ ), rata-rata lama sekolah ( $X_7$ ) dan tingkat pengangguran terbuka ( $X_{10}$ ). Selain dari variabel angka melek huruf ( $X_6$ ) dan rumah tangga dengan sumber penerangan PLN ( $X_8$ ) untuk variabel jumlah penduduk miskin ( $X_9$ ) memiliki nilai rata-rata yang paling rendah dibandingkan klaster yang lainnya..

#### E. Daftar Pustaka

- [1] B. RI, "Indikator Kesejahteraan Rakyat 2023," 2023. <https://bps.go.id> (accessed May 05, 2024).
- [2] N. BPS, Indikator Kesejahteraan Rakyat Provinsi Nusa Tenggara Timur 2022. Kupang: BPS Provinsi Nusa Tenggara Timur, 2022.



- [3] N. BPS, Provinsi Nusa Tenggara Timur Dalam Angka 2023. Kupang: BPS Provinsi Nusa Tenggara Timur, 2023.
- [4] J. F. Hair Jr., W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis Seventh Edition*. New Jersey: Pearson Prentice Hall, 2010.
- [5] R. M. R. Cesario, W. Syafmen, and B. Mardhotillah, "Analisis Cluster Metode Average Linkage Berdasarkan Indikator Pendidikan tiap Provinsi Di Indonesia," *Median. Ilm. Pop.*, vol. 5, no. 1, pp. 30–39, 2022.
- [6] V. A. Latuhimallo, M. W. Talakua, and Z. A. Leleury, "Analisis Clustering untuk Pengelompokan Kabupaten/Kota Berdasarkan Indikator Kesejahteraan Rakyat di Wilayah Provinsi Maluku," *Param. J. Mat. Stat. dan Ter.*, vol. 1, no. 1, pp. 01–12, 2023.
- [7] N. W. R. Damayanthi, N. L. P. Suciptawati, K. Jayanegara, and E. N. Kencana, "Pengelompokan Provinsi di Indonesia Menurut Indikator Indeks Kebahagiaan Menggunakan Metode Average Linkage," *Innov. J. Soc. Sci. Res.*, vol. 3, no. 4, pp. 8859–8872, 2023.
- [8] Y. Reinaldi, N. Ulinuha, and M. Hafiyusholeh, "Comparison of Single Linkage, Complete Linkage, and Average Linkage Methods on Community Welfare Analysis in Cities and Regencies in East Java," *J. Mat. Stat. dan Komputasi*, vol. 18, no. 1, pp. 130–140, 2021, doi: 10.20956/j.v18i1.14228.
- [9] R. Ramadani and A. Salma, "Metode Average Linkage dan Ward dalam Pengelompokan Kesejahteraan Sumatera Barat Tahun 2021," *J. Math. UNP*, vol. 7, no. 3, pp. 11–24, 2022.
- [10] S. Santoso, *Statistik Multivariat dengan SPSS*. Jakarta: PT Elex Media Komputindo, 2017.
- [11] N. S. P. Nasution, "Uji Manual KMO, Barlet, MSA," 2018. <https://www.scrib.com> (accessed Jan. 29, 2024).
- [12] M. Hasrul, "Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Provinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Rakyat," Universitas Islam Negeri (UIN) Alauddin Makassar Oleh, 2018.
- [13] C. E. Mongi, "Penggunaan Analisis Two Step Clustering untuk Data Campuran," *JdC*, vol. 4, no. 1, pp. 9–19, 2015, doi: 10.35799/dc.4.1.2015.7251.
- [14] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis Sixth Edition*. New Jersey: Pearson Prentice Hall, 2007.
- [15] S. Sunaryo, Setiawan, and T. H. Siagian, "Mengatasi Masalah Multikolinearitas dan Outlier Dengan Pendekatan ROBPCA (Studi Kasus Analisis Regresi Angka Kematian Bayi di Jawa Timur)," *J. Mat. Sains dan Teknol.*, vol. 12, no. 1, pp. 1–10, 2011, [Online]. Available: <http://www.wis.kuleuven.ac.be/stat/robust.html>.