



## **SKRIPSI**

# **SEGMENTASI WILAYAH INDONESIA BERDASARKAN PENGELOLAAN SAMPAH TAHUN 2023 MENGGUNAKAN *MEAN SHIFT***

**AWAL LIDYA MUSAFFAK**  
NPM 21083010088

**DOSEN PEMBIMBING**  
Kartika Maulida Hindrayani, S.Kom., M.Kom.  
Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI SAINS DATA  
SURABAYA  
2025**



## **SKRIPSI**

# **SEGMENTASI WILAYAH INDONESIA BERDASARKAN PENGELOLAAN SAMPAH TAHUN 2023 MENGGUNAKAN *MEAN SHIFT***

**AWAL LIDYA MUSAFFAK**  
NPM 21083010088

**DOSEN PEMBIMBING**  
Kartika Maulida Hindrayani, S.Kom., M.Kom.  
Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI SAINS DATA  
SURABAYA  
2025**



**SKRIPSI**

**SEGMENTASI WILAYAH INDONESIA  
BERDASARKAN PENGELOLAAN SAMPAH  
MENGUNAKAN *MEAN SHIFT***

**AWAL LIDYA MUSAFFAK**  
NPM 21083010088

**DOSEN PEMBIMBING**  
Kartika Maulida Hindrayani, S.Kom., M.Kom.  
Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI  
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI SAINS DATA  
SURABAYA  
2025**

**LEMBAR PENGESAHAN**

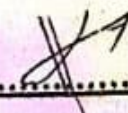
**SEGMENTASI WILAYAH INDONESIA BERDASARKAN  
PENGELOLAAN SAMPAH TAHUN 2023 MENGGUNAKAN MEAN  
SHIFT**

Oleh :  
**AWAL LIDYA MUSAFFAK**  
NPM. 21083010088

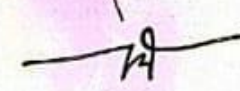
Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Sidang Skripsi Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 5 Juni 2025:

**Menyetujui,**

**Kartika Maulida Hindrayani, S.Kom., M.Kom.**  
NIP. 19920909 202203 2 009

  
..... (Pembimbing I)

**Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.**  
NIP. 19830310 202121 1 006

  
..... (Pembimbing II)

**Trimono, S.Si., M.Si.**  
NIP. 19950908 202203 1 003

  
..... (Ketua Penguji)

**Alfan Rizaldy Pratama, S.Tr.T., M.Tr.Kom.**  
NIP. 19990606 202406 1 001

  
..... (Penguji I)

**Mengetahui,**  
**Dekan Fakultas Ilmu Komputer**

  
  
**Prof. Dr. Ir. Novirina Hendrasarie, MT.**  
NIP. 19681126 199403 2 001

**LEMBAR PERSETUJUAN**

**SEGMENTASI WILAYAH INDONESIA BERDASARKAN  
PENGELOLAAN SAMPAH TAHUN 2023 MENGGUNAKAN *MEAN  
SHIFT***

Oleh:  
**AWAL LIDYA MUSAFFAK**  
NPM. 21083010088

Telah disetujui untuk mengikuti Ujian Skripsi



Menyetujui,

**Koordinator Program Studi Sains Data  
Fakultas Ilmu Komputer**

**Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.**  
NIP. 19801205 200501 1 002

## SURAT PERNYATAAN BEBAS PLAGIASI

Yang bertandatangan di bawah ini:

Nama : Awal Lidya Musaffak  
NPM : 21083010088  
Program : Sarjana (S1)  
Program Studi : Sains Data  
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apalagi dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 16 Juni 2025  
Yang Membuat Pernyataan,



Awal Lidya Musaffak  
NPM. 21083010088

## ABSTRAK

Nama Mahasiswa / NPM : Awal Lidya Musaffak / 21083010088  
Judul Skripsi : Segmentasi Wilayah Indonesia Berdasarkan Pengelolaan Sampah Tahun 2023 Menggunakan *Mean Shift*  
Dosen Pembimbing : 1. Kartika Maulida Hindrayani, S.Kom., M.Kom.  
2. Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.

Sampah merupakan bagian yang tidak terpisahkan dari kehidupan manusia, dan volumenya terus meningkat seiring pertumbuhan jumlah penduduk. Berdasarkan data dari Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) Kementerian Lingkungan Hidup dan Kehutanan (KLHK), pada tahun 2023 Indonesia menghasilkan timbulan sampah harian sebesar 106.145,71 ton dan timbulan sampah tahunan mencapai 38.743.185,18 ton. Jumlah ini mencerminkan masih lemahnya sistem pengelolaan sampah, terutama dalam penanganan dan pengurangan sampah rumah tangga serta sejenis rumah tangga. Pemerintah Indonesia telah melakukan berbagai upaya melalui sejumlah program dan kebijakan, salah satunya adalah Kebijakan dan Strategi Nasional (Jakstranas) yang menjadi pedoman pengelolaan sampah secara nasional. Jakstranas menetapkan target pengurangan dan penanganan sampah, namun hingga tahun 2023, target tersebut belum tercapai. Oleh karena itu, diperlukan pendekatan berbasis data untuk mengelompokkan wilayah berdasarkan capaian pengurangan dan penanganan sampah. Metode klustering menjadi pilihan yang relevan karena dapat mengidentifikasi wilayah dengan karakteristik pengelolaan sampah yang serupa. Hasil pengelompokan ini dapat digunakan sebagai dasar dalam menetapkan prioritas kebijakan dan mengalokasikan sumber daya secara lebih tepat dan efisien. Dalam penelitian ini, algoritma *Mean Shift* dipilih karena memiliki keunggulan dalam menentukan jumlah kluster secara otomatis tanpa memerlukan *input* jumlah kluster di awal. Selain itu, *Mean Shift* mampu menangani bentuk kluster yang tidak simetris dan beragam, yang sesuai dengan karakteristik data capaian pengelolaan sampah antar wilayah yang kompleks dan tidak seragam. Pada penelitian ini didapatkan bahwa metode *Mean Shift* yang terbaik dengan menghasilkan nilai DBI sebesar 0,600 dan *silhouette* sebesar 0,649. Hasil yang didapat kluster 0 sebagai kluster tingkat pengelolaan sampah rendah sedangkan kluster 1 tinggi.

**Kata kunci :** *Davies Bouldin Index*, Klustering, *Mean Shift*, Pengelolaan sampah, *Silhouette score*

*Halaman ini sengaja dikosongkan*

## **ABSTRACT**

*Student Name / NPM* : Awal Lidya Musaffak / 21083010088  
*Thesis Title* : Indonesian Regional Segmentation Based on Waste Management in 2023 Using Mean Shift  
*Advisor* : 1. Kartika Maulida Hindrayani, S.Kom., M.Kom.  
2. Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.

## **ABSTRACT**

*Waste is an inseparable part of human life, and its volume continues to increase in line with population growth. According to data from the National Waste Management Information System (SIPSN) of the Ministry of Environment and Forestry (KLHK), in 2023, Indonesia generated a daily waste volume of 106,145.71 tons and an annual total of 38,743,185.18 tons. These figures reflect the continued weakness of the waste management system, particularly in reducing and handling household and household-like waste. The Indonesian government has implemented various programs and policies to address this issue, one of which is the National Policy and Strategy (Jakstranas), which serves as a national guideline for waste management. Jakstranas set targets for waste reduction and handling; however, as of 2023, these targets had not been achieved. Therefore, a data-driven approach is needed to group regions based on their waste reduction and handling performance. Clustering methods are a relevant approach because they can identify regions with similar waste management characteristics. The results of this clustering can serve as a basis for setting policy priorities and allocating resources more accurately and efficiently. This study adopts the Mean Shift algorithm due to its ability to automatically determine the number of clusters without requiring a predefined input. Moreover, Mean Shift is capable of identifying non-symmetric and varied cluster shapes, which aligns well with the complex and non-uniform nature of waste management data across regions. In this study, Mean Shift yielded the best results, with a DBI of 0,600 and a Silhouette of 0,649. The clustering results showed that Cluster 0 represents regions with low waste management performance, while Cluster 1 indicates regions with high performance.*

**Keywords:** *Clustering, Davies Bouldin Index, Mean Shift, Silhouette score, Waste management*

*Halaman ini sengaja dikosongkan*

## KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Segmentasi Wilayah Indonesia Berdasarkan Pengelolaan Sampah Tahun 2023 Menggunakan *Mean Shift*”** dapat terselesaikan dengan baik.

Penyusunan skripsi ini tidak terlepas dari berbagai tantangan. Namun, berkat dukungan dan bantuan dari berbagai pihak, skripsi ini dapat diselesaikan dengan baik. Untuk itu penulis mengucapkan terima kasih kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Bapak Dr. Eng. Ir. Dwi Arman Prasetya., ST., MT., IPU., Asean, Eng, selaku Ketua Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran “ Jawa Timur.
3. Ibu Kartika Maulida Hindrayani, S.Kom., M.Kom. selaku Dosen Pembimbing I, atas bimbingan , arahan, dan kesabaran dalam memberikan dukungan selama proses penyusunan skripsi ini.
4. Bapak Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T. selaku Dosen Pembimbing II, atas bimbingan, masukan, dan dukungan dalam penyelesaian skripsi ini.
5. Bapak dan Ibu Dosen Program Studi Sains Data yang telah membimbing serta memberikan arahan selama masa perkuliahan.
6. Kedua orang tua, adik, dan keluarga yang senantiasa mendoakan dan memberikan dukungan kepada penulis.
7. Teman-teman yang kebersamaan penulis selama proses penyusunan skripsi berlangsung yang selalu memberikan masukan, dukungan, dan motivasi untuk mampu menyelesaikan skripsi ini.
8. Seluruh pihak yang terlibat dalam penyusunan skripsi ini, yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, Juni 2025

Penulis

## DAFTAR ISI

|                                                    |             |
|----------------------------------------------------|-------------|
| <b>HALAMAN SAMPUL</b> .....                        | <b>i</b>    |
| <b>LEMBAR PENGESAHAN</b> .....                     | <b>ii</b>   |
| <b>LEMBAR PERSETUJUAN</b> .....                    | <b>iii</b>  |
| <b>SURAT PERNYATAAN BEBAS PLAGIASI</b> .....       | <b>iv</b>   |
| <b>ABSTRAK</b> .....                               | <b>v</b>    |
| <b>ABSTRACT</b> .....                              | <b>vii</b>  |
| <b>KATA PENGANTAR</b> .....                        | <b>ix</b>   |
| <b>DAFTAR ISI</b> .....                            | <b>xi</b>   |
| <b>DAFTAR GAMBAR</b> .....                         | <b>xiii</b> |
| <b>DAFTAR TABEL</b> .....                          | <b>xv</b>   |
| <b>DAFTAR LAMPIRAN</b> .....                       | <b>xvii</b> |
| <b>DAFTAR NOTASI</b> .....                         | <b>xix</b>  |
| <b>BAB I PENDAHULUAN</b> .....                     | <b>1</b>    |
| 1.1. Latar Belakang .....                          | 1           |
| 1.2. Rumusan Masalah .....                         | 6           |
| 1.3. Batasan Masalah.....                          | 6           |
| 1.4. Tujuan Penelitian.....                        | 6           |
| 1.5. Manfaat Penelitian.....                       | 7           |
| <b>BAB II TINJAUAN PUSTAKA</b> .....               | <b>9</b>    |
| 2.1. Penelitian Terdahulu .....                    | 9           |
| 2.2. Dasar Teori.....                              | 15          |
| 2.2.1. Pengelolaan Sampah .....                    | 15          |
| 2.2.2. SIPSN .....                                 | 19          |
| 2.2.3. Klastering .....                            | 20          |
| 2.2.4. <i>Kernel Density Estimation</i> (KDE)..... | 20          |
| 2.2.5. <i>Mean Shift</i> .....                     | 21          |
| 2.2.6. <i>Metode Elbow</i> .....                   | 24          |
| 2.2.7. <i>K-Means</i> .....                        | 25          |
| 2.2.8. <i>Davies Bouldin Index</i> (DBI).....      | 27          |
| 2.2.9. <i>Silhouette Score</i> .....               | 28          |
| 2.2.10. Standarisasi Data.....                     | 29          |

|                                                    |           |
|----------------------------------------------------|-----------|
| <b>BAB III DESAIN DAN IMPLEMENTASI SISTEM.....</b> | <b>31</b> |
| 3.1. Variabel Penelitian dan Sumber Data .....     | 31        |
| 3.2. Langkah Analisis .....                        | 31        |
| 3.3. Desain Sistem .....                           | 40        |
| <b>BAB IV HASIL DAN PEMBAHASAN .....</b>           | <b>45</b> |
| 4.1. Analisis Deskriptif .....                     | 45        |
| 4.2. <i>Preprocessing</i> .....                    | 46        |
| 4.2.1. Menghapus Variabel .....                    | 46        |
| 4.2.2. Mengubah Nama Variabel .....                | 47        |
| 4.2.3. <i>Missing Value</i> .....                  | 48        |
| 4.2.4. Deteksi Data Duplikat .....                 | 50        |
| 4.2.5. Deteksi <i>Outlier</i> .....                | 50        |
| 4.2.6. Pemilihan Variabel .....                    | 54        |
| 4.2.7. Standarisasi Data .....                     | 55        |
| 4.2.8. Statistika Deskriptif .....                 | 59        |
| 4.2.9. Visualisasi Data .....                      | 60        |
| 4.3. Pemodelan .....                               | 61        |
| 4.3.1. <i>Mean Shift</i> .....                     | 61        |
| 4.3.1. <i>K-Means</i> .....                        | 67        |
| 4.3.2. Pemilihan Model Terbaik .....               | 71        |
| 4.4. <i>Deployment</i> .....                       | 73        |
| <b>BAB V PENUTUP .....</b>                         | <b>79</b> |
| 5.1. Kesimpulan .....                              | 79        |
| 5.2. Saran Pengembangan .....                      | 80        |
| <b>DAFTAR PUSTAKA .....</b>                        | <b>83</b> |
| <b>LAMPIRAN .....</b>                              | <b>89</b> |

## DAFTAR GAMBAR

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| <b>Gambar 2.1</b> Proses Pengelolaan Sampah .....                        | 16 |
| <b>Gambar 2.2</b> Logo SIPSN .....                                       | 19 |
| <b>Gambar 2.3</b> Klastering.....                                        | 20 |
| <b>Gambar 2.4</b> Inisialisasi Titik Awal <i>Mean Shift</i> .....        | 22 |
| <b>Gambar 2.5</b> Pencarian Titik Pusat Baru <i>Mean Shift</i> .....     | 23 |
| <b>Gambar 2.6</b> Titik Konvergensi <i>Mean Shift</i> .....              | 24 |
| <b>Gambar 2.7</b> Inisialisasi <i>Centroid K-Means</i> .....             | 25 |
| <b>Gambar 2.8</b> Menghitung Jarak Data ke <i>Centroid K-Means</i> ..... | 26 |
| <b>Gambar 3.1</b> Diagram Alir Penelitian.....                           | 33 |
| <b>Gambar 3.2</b> <i>Use Case Streamlit</i> .....                        | 42 |
| <b>Gambar 3.3</b> Desain Tampilan Menu.....                              | 43 |
| <b>Gambar 3.4</b> Desain Tampilan <i>Upload Data</i> .....               | 43 |
| <b>Gambar 3.5</b> Desain Tampilan Pemodelan .....                        | 44 |
| <b>Gambar 3.6</b> Desain Tampilan Visualisasi .....                      | 44 |
| <b>Gambar 4.1</b> Mendeteksi Data Duplikat .....                         | 50 |
| <b>Gambar 4.2</b> Deteksi <i>outlier</i> .....                           | 52 |
| <b>Gambar 4.3</b> <i>Handling outlier</i> .....                          | 54 |
| <b>Gambar 4.4</b> Histogram Variabel.....                                | 60 |
| <b>Gambar 4.5</b> <i>Heatmap</i> Korelasi .....                          | 61 |
| <b>Gambar 4.6</b> Perbandingan <i>Bandwidth</i> .....                    | 66 |
| <b>Gambar 4.7</b> Grafik Metode <i>Elbow</i> .....                       | 68 |
| <b>Gambar 4.8</b> Visualisasi <i>Mean Shift</i> .....                    | 71 |
| <b>Gambar 4.9</b> Rata-rata Klaster .....                                | 72 |
| <b>Gambar 4.10</b> Menu Utama .....                                      | 73 |
| <b>Gambar 4.11</b> Menu <i>Upload Data</i> .....                         | 74 |
| <b>Gambar 4.12</b> Menu Pemodelan .....                                  | 74 |
| <b>Gambar 4.13</b> Hasil Menu Pemodelan.....                             | 75 |
| <b>Gambar 4.14</b> Visualisasi Klaster 0 .....                           | 75 |
| <b>Gambar 4.15</b> Visualisasi Diagram Klaster 0 .....                   | 76 |

|                                                       |    |
|-------------------------------------------------------|----|
| <b>Gambar 4.16</b> Visualisasi Tabel Klaster 0.....   | 76 |
| <b>Gambar 4.17</b> Visualisasi Klaster 1.....         | 77 |
| <b>Gambar 4.18</b> Visualisasi Diagram Klaster 1..... | 77 |
| <b>Gambar 4.19</b> Visualisasi Tabel Klaster 1.....   | 78 |

## DAFTAR TABEL

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| <b>Tabel 3.1</b> Tabel Timbulan Sampah .....                             | 31 |
| <b>Tabel 3.2</b> Tabel Capaian Kinerja Pengelolaan Sampah .....          | 31 |
| <b>Tabel 3.3</b> Tabel Lanjutan Capaian Kinerja Pengelolaan Sampah ..... | 32 |
| <b>Tabel 3.4</b> Variabel Yang Digunakan .....                           | 32 |
| <b>Tabel 4.1</b> Tabel <i>Merge</i> Data .....                           | 45 |
| <b>Tabel 4.2</b> Tabel <i>Merge</i> Data .....                           | 46 |
| <b>Tabel 4.3</b> Hasil Penghapusan Variabel.....                         | 47 |
| <b>Tabel 4.4</b> Perubahan Nama Variabel.....                            | 47 |
| <b>Tabel 4.5</b> Tabel <i>Missing Value</i> .....                        | 48 |
| <b>Tabel 4.6</b> Tabel Penanganan <i>Missing Value</i> .....             | 49 |
| <b>Tabel 4.7</b> Lima Data Teratas .....                                 | 56 |
| <b>Tabel 4.8.</b> Hasil Standarisasi Data.....                           | 58 |
| <b>Tabel 4.9</b> Statistika Deskriptif .....                             | 59 |
| <b>Tabel 4.10</b> Hasil Titik Pusat <i>Mean Shift</i> .....              | 64 |
| <b>Tabel 4.11</b> Tabel Evaluasi <i>Mean Shift</i> .....                 | 66 |
| <b>Tabel 4.12</b> Hasil Pengelompokan <i>K-Means</i> .....               | 70 |
| <b>Tabel 4.13</b> Klaster wilayah.....                                   | 72 |

*Halaman ini sengaja dikosongkan*

## DAFTAR LAMPIRAN

|                                                                        |    |
|------------------------------------------------------------------------|----|
| <b>Lampiran 1.</b> Data Penelitian.....                                | 89 |
| <b>Lampiran 2.</b> <i>Script</i> Analisis.....                         | 90 |
| <b>Lampiran 3.</b> <i>Script</i> Pembuatan <i>User Interface</i> ..... | 91 |
| <b>Lampiran 4.</b> LoA .....                                           | 92 |

*Halaman ini sengaja dikosongkan*

## DAFTAR NOTASI

|               |                                                                                               |
|---------------|-----------------------------------------------------------------------------------------------|
| $\hat{f}(x)$  | : Estimasi fungsi kerapatan pada titik $x$                                                    |
| $N$           | : Jumlah total data                                                                           |
| $h$           | : <i>Bandwidth</i>                                                                            |
| $D$           | : Dimensi dari data                                                                           |
| $K$           | : <i>Kernel function</i>                                                                      |
| $x_i$         | : Data pengamatan ke- $i$                                                                     |
| $m(x)$        | : Titik pusat baru                                                                            |
| $x$           | : Titik pusat saat ini                                                                        |
| $c_k$         | : Pusat klaster ke $k$                                                                        |
| $D(x_i, x_j)$ | : Jarak data ke $i$ ke pusat klaster $j$                                                      |
| $x_{ki}$      | : Data ke $i$ pada atribut ke $k$                                                             |
| $x_{kj}$      | : Data ke $j$ pada atribut ke $k$                                                             |
| $C_{m(q)}$    | : <i>Centroid</i> dari klaster $m$ pada perhitungan ke $q$                                    |
| $N_m$         | : Banyak data pada klaster $m$                                                                |
| $x_{i(q)}$    | : Data ke $i$ pada iterasi ke $q$                                                             |
| $m_i$         | : Jumlah data dalam klaster ke- $i$                                                           |
| $C_j$         | : <i>Centroid</i> klaster ke- $j$                                                             |
| $d(x_j, C_j)$ | : Jarak setiap data ke <i>centroid</i> $i$ yang dihitung menggunakan jarak <i>euclidiance</i> |
| $SSW_i$       | : <i>Sum of Square Within Cluster</i> pada <i>centroid</i> $i$                                |
| $SSB_{i,j}$   | : <i>Sum of Square Between Cluster</i> data ke $i$ dengan $j$ pada klaster yang berbeda       |
| $R_{i,j}$     | : Rasio dari nilai SSW dan SSB                                                                |
| $S$           | : <i>Silhouette</i>                                                                           |
| $a$           | : Jarak rata-rata antara sampel dengan semua titik lain di klaster yang sama                  |
| $b$           | : Jarak rata-rata antara sampel dengan semua titik lain di klaster terdekat                   |

*Halaman ini sengaja dikosongkan*