

## Analisis *Cluster* Hierarki Metode *Single Linkage* Pada Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Indikator Kesehatan

Diah Rahmahadi Thamrin<sup>1</sup>, Dewi Murni<sup>2</sup>

<sup>1,2</sup>Prodi Matematika, Fakultas Matematika Ilmu Pengetahuan dan Alam Universitas Negeri Padang (UNP)

---

### Article Info

#### Article history:

Received May 27, 2022

Revised August 03, 2022

Accepted September 15, 2022

---

#### Keywords:

Hierarchical Cluster Analysis  
Single Linkage  
Health Indicator

#### Kata Kunci:

Analisis *Cluster* Hierarki  
*Single Linkage*  
Indikator Kesehatan

### ABSTRACT

Health indicators are variables that are needed to estimate the conditions and changes faced by an area in the health sector. Health indicators are needed by the government to plan development in the health sector in order to achieve optimal results in accordance with the conditions of each region. This study aims to group districts/cities in West Sumatra Province based on health indicators in 2020 and interpret the characteristics of each cluster that has been formed by using a single linkage method of hierarchical cluster analysis. The results of the grouping are divided into five clusters, the first cluster consists of the Mentawai Islands Regency. The second cluster consists of 12 regencies/cities namely Pesisir Selatan Regency, Padang Pariaman Regency, West Pasaman Regency, Agam Regency, Pasaman Regency, Dharmasraya Regency, South Solok Regency, Sawahlunto City, Bukittinggi City, Payakumbuh City, Solok City and Padang Panjang City. The third cluster consists of 4 regencies, namely Solok Regency, Tanah Datar Regency, Fifty Kota Regency and Sijunjung Regency. The fourth cluster consists of Kota Pariaman. The fifth cluster consists of the City of Padang.

### ABSTRAK

Indikator kesehatan merupakan variabel yang diperlukan untuk memperkirakan kondisi dan perubahan yang dihadapi suatu daerah di bidang kesehatan. Indikator kesehatan sangat diperlukan oleh pemerintah untuk merencanakan pembangunan dibidang kesehatan agar mencapai hasil yang optimal sesuai dengan kondisi daerah masing-masing. Penelitian ini bertujuan untuk mengelompokkan kabupaten/kota di Provinsi Sumatera Barat berdasarkan indikator kesehatan tahun 2020 dan menginterpretasikan karakteristik pada setiap *cluster* yang telah terbentuk dengan analisis *cluster* hierarki metode *single linkage*. Hasil pengelompokannya terbagi menjadi lima *cluster*, *Cluster* pertama terdiri dari Kabupaten Kepulauan Mentawai. *Cluster* kedua terdiri dari 12 kabupaten/kota yaitu Kabupaten Pesisir Selatan, Kabupaten Padang Pariaman, Kabupaten Pasaman Barat, Kabupaten Agam, Kabupaten Pasaman, Kabupaten Dharmasraya, Kabupaten Solok Selatan, Kota Sawahlunto, Kota Bukittinggi, Kota Payakumbuh, Kota Solok dan Kota Padang Panjang. *Cluster* ketiga terdiri dari 4 kabupaten yaitu Kabupaten Solok, Kabupaten Tanah Datar, Kabupaten Lima Puluh Kota dan Kabupaten Sijunjung. *Cluster* keempat terdiri dari Kota Pariaman. *Cluster* kelima terdiri dari Kota Padang.

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



---

### Penulis pertama

(Diah Rahmahadi Thamrin)

Prodi Matematika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Negeri Padang, Jl. Prof. Dr. Hamka, Air Tawar barat, Padang Utara, Padang, 25171  
Email: [diah.rahmahadi04@gmail.com](mailto:diah.rahmahadi04@gmail.com)

Padang, Sumatera Barat

## 1. PENDAHULUAN

Keberhasilan terhadap program-program yang dibangun dalam bidang kesehatan yang telah dilakukan oleh pemerintah dapat dipertimbangkan dengan menggunakan indikator kesehatan [1]. Indikator kesehatan merupakan variabel yang diperlukan untuk memperkirakan kondisi dan perubahan yang dihadapi suatu daerah di bidang kesehatan [2]. Indikator kesehatan sangat diperlukan oleh pemerintah untuk merencanakan pembangunan dibidang kesehatan agar mencapai hasil yang optimal sesuai dengan kondisi daerah masing-masing.

Masih kurangnya pengetahuan tentang kesehatan membuat seseorang memiliki pola hidup yang tidak sehat sehingga lama kelamaan akan mengalami gangguan kesehatan yang disebut sebagai morbiditas atau kesakitan. Morbiditas adalah keluhan kesehatan yang dialami oleh penduduk sehingga dapat menyebabkan aktivitas sehari-hari terganggu seperti saat sekolah, bekerja atau kegiatan lainnya sebagaimana biasanya. Morbiditas yang tinggi artinya semakin meningkatnya keluhan kesehatan pada penduduk. Hal tersebut dapat menunjukkan rendahnya tingkat kesehatan masyarakat.

Di Provinsi Sumatera Barat pada tahun 2020, angka morbiditas penduduknya adalah 14,39% yang lebih besar dari tahun 2019 yaitu sebesar 13,54%. Hal ini menunjukkan bahwa tingkat kesehatan penduduk di Provinsi Sumatera Barat menurun dalam satu tahun terakhir. Oleh sebab itu permasalahan kesehatan tersebut harus diselesaikan karena akan berdampak pada tingginya angka kesakitan yang menunjukkan rendahnya tingkat kesehatan penduduk di Provinsi Sumatera Barat yang nantinya akan berdampak juga pada kondisi tubuh seseorang, sehingga dapat mengganggu aktivitas atau kegiatan lainnya sebagaimana biasanya dikarenakan mengalami gangguan atau keluhan kesehatan. Maka dari itu, diperlukan pengelompokan daerah kabupaten/ kota di Provinsi Sumatera Barat untuk memudahkan proses penanganan masalah kesehatan tersebut. Dalam mengelompokkan daerah kabupaten/kota berdasarkan kemiripan indikator kesehatan di Sumatera Barat digunakan analisis statistika yaitu analisis *cluster* hierarki dengan metode *single linkage*.

Analisis *cluster* merupakan pengelompokan sejumlah  $n$  objek berdasarkan  $p$  variabel yang memiliki kesamaan ciri khusus diantara objek-objek tersebut, sehingga keragaman dalam suatu *cluster* tersebut lebih kecil dibandingkan dengan keragaman antar *cluster*. Misalkan terdapat  $n$  objek dan  $p$  variabel maka observasinya  $x_{ij}$  dengan  $i = 1, 2, \dots, n$  dan  $j = 1, 2, \dots, p$  [3]. Metode *single linkage* jarak dua *cluster* dapat dihitung dengan jarak minimum atau jarak terdekat (*nearest neighbor*) antara sebuah objek dalam *cluster* yang satu dengan sebuah objek dalam *cluster* yang lain. Langkah pertama kali yang harus dilakukan pada metode *single linkage* ini adalah harus memilih jarak terdekat antara *cluster-cluster* tersebut yang dapat dirumuskan dengan  $D = \{d_{ik}\}$ . Kemudian digabungkan objek-objek tersebut dengan melihat jarak yang terdekat, misalnya objek tersebut disimbolkan dengan *cluster*  $u$  dan *cluster*  $v$  untuk memperoleh *cluster* gabungan ( $uv$ ). Untuk menghitung jarak *cluster* ( $uv$ ) ke *cluster* lainnya dapat dirumuskan sebagai berikut [4] :

$$d_{(uv)w} = \min(d_{uw}, d_{vw})$$

Keterangan:

$d_{uw}$  adalah jarak terdekat dari *cluster*  $u$  dan  $w$

$d_{vw}$  adalah jarak terdekat dari *cluster*  $v$  dan  $w$

dan menggunakan jarak manhattan dengan rumus [5]:

$$d_{i,j} = \sum_{k=1}^n |x_{i,k} - x_{j,k}|$$

Keterangan:

$d_{i,j}$  adalah jarak antara objek ke- $i$  dengan objek ke- $j$

$x_{i,k}$  adalah nilai objek ke- $i$  pada variabel ke- $k$

$x_{j,k}$  adalah nilai objek ke- $j$  pada variabel ke- $k$

$n$  adalah banyaknya variabel



## 2. METODE

Jenis penelitian ini adalah terapan dengan menggunakan data sekunder. Data bersumber dari buku publikasi Badan Pusat Statistik (BPS) Provinsi Sumatera Barat berupa data Indikator Kesehatan 2020 dan variabel yang digunakan terdiri dari jumlah penduduk yang mempunyai keluhan kesehatan ( $X_1$ ), jumlah penduduk yang mempunyai keluhan kesehatan dan berobat jalan ( $X_2$ ), jumlah penduduk yang pernah rawat inap ( $X_3$ ), jumlah penduduk yang menggunakan jaminan kesehatan untuk rawat inap ( $X_4$ ), dan rata-rata lama rawat inap ( $X_5$ ).

Langkah-langkah analisis data pada penelitian ini yaitu

- a. Mengumpulkan data yang akan diolah
- b. Melakukan standarisasi pada data yang didapat apabila data tersebut memiliki perbedaan satuan dengan *Z-Score*.
- c. Menggunakan uji koefisien korelasi *pearson* untuk mengetahui adanya hubungan atau tidak antar variabel.
- d. Menggunakan jarak manhattan atau *city-block* untuk menghitung jarak tiap variabel.
- e. Membentuk data kabupaten atau kota di Provinsi Sumatera Barat menjadi *cluster-cluster* menggunakan metode *single linkage*.
- f. Melakukan pemotongan dendrogram yang dihasilkan dari analisis *cluster* hierarki metode *single linkage* berdasarkan selisih jarak penggabungan terbesar.
- g. Menghitung rata-rata variabel setiap kelompok atau *cluster*.
- h. Menginterpretasikan hasil yang sudah didapat dan menarik kesimpulan.

## 3. HASIL DAN PEMBAHASAN

Data yang digunakan sebagai bahan penelitian berupa tahun 2020 di Provinsi Sumatera Barat yang terdiri dari 12 kabupaten dan 7 kota. Sebelum data dianalisis, terlebih dahulu mendeskripsikan variabel yang digunakan pada data indikator kesehatan di Sumatera Barat.

Tabel 1. Statistika Deskriptif Indikator Kesehatan di Provinsi Sumatera Barat

Variabel	N	Mean	Maximum	Minimum	Std. Deviation
$X_1$	19	91.550,53	302.357	18.231	70.579,8
$X_2$	19	52.299,53	188.882	7.915	43.978,723
$X_3$	19	14.890,68	59.836	1.141	12.817,95
$X_4$	19	11.384,95	50.927	830	10.732,74
$X_5$	19	4,8874	6,15	3,88	0,57333

Berdasarkan Tabel 1, dapat diamati bahwa rata-rata jumlah penduduk yang mempunyai keluhan kesehatan adalah 91.550,53 orang. Dengan jumlah tertinggi sebesar 302.357 orang dan yang terendah adalah 18.231 orang dengan standar deviasi sebesar 70.579,8 orang.

### 3.1. Uji Korelasi

Setelah data dideskripsikan dan distandarisasi maka selanjutnya melakukan uji korelasi antar variabel. Sebelum analisis *cluster* dilakukan, terdapat asumsi yang harus dipenuhi yaitu tidak adanya hubungan antar variabel atau tidak terjadi multikolinearitas. Pada uji korelasi ini dilakukan perhitungan nilai koefisien korelasi *pearson* ( $r$ ) dengan taraf signifikan 5%. Dengan kriteria uji,

(Diah Rahmahadi Thamrin)

jika  $-r_{tabel} \leq r_{hitung} \leq +r_{tabel}$  maka  $H_0$  diterima yang berarti tidak ada hubungan antar 2 variabel dan jika  $r_{hitung} \geq r_{tabel}$  maka  $H_0$  ditolak yang berarti terdapat hubungan antar 2 variabel. Untuk nilai koefisien dan signifikan korelasi *pearson* antar variabel dapat diamati pada Tabel 2 serta ringkasan uji korelasi dapat diamati pada Tabel 3.

Tabel 2. Nilai Koefisien dan Signifikan Korelasi *Pearson*

	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
X <sub>1</sub>	1	0,990	0,950	0,916	-0,146
	0	0	0	0	0,552
X <sub>2</sub>	0,990	1	0,972	0,944	-0,139
	0	0	0	0	0,570
X <sub>3</sub>	0,950	0,972	1	0,991	-0,103
	0	0	0	0	0,676
X <sub>4</sub>	0,916	0,944	0,991	1	-0,084
	0	0	0	0	0,731
X <sub>5</sub>	-0,146	-0,139	-0,103	-0,084	1
	0,552	0,570	0,676	0,731	0

Tabel 3. Ringkasan Uji Korelasi *Pearson*

Variabel	$r_{hitung}$	$r_{tabel}$	Keputusan	Kesimpulan
X <sub>1</sub> dan X <sub>2</sub>	0,990	0,4555	0,990 > 0,4555	$H_0$ ditolak
X <sub>1</sub> dan X <sub>3</sub>	0,950	0,4555	0,950 > 0,4555	$H_0$ ditolak
X <sub>1</sub> dan X <sub>4</sub>	0,916	0,4555	0,916 > 0,4555	$H_0$ ditolak
X <sub>1</sub> dan X <sub>5</sub>	-0,146	0,4555	$-0,4555 < -0,146 < 0,4555$	$H_0$ diterima
X <sub>2</sub> dan X <sub>3</sub>	0,972	0,4555	0,972 > 0,4555	$H_0$ ditolak
X <sub>2</sub> dan X <sub>4</sub>	0,944	0,4555	0,944 > 0,4555	$H_0$ ditolak
X <sub>2</sub> dan X <sub>5</sub>	-0,139	0,4555	$-0,4555 < -0,139 < 0,4555$	$H_0$ diterima
X <sub>3</sub> dan X <sub>4</sub>	0,991	0,4555	0,991 > 0,4555	$H_0$ ditolak
X <sub>3</sub> dan X <sub>5</sub>	-0,103	0,4555	$-0,4555 < -0,103 < 0,4555$	$H_0$ diterima
X <sub>4</sub> dan X <sub>5</sub>	-0,084	0,4555	$-0,4555 < -0,084 < 0,4555$	$H_0$ diterima

Berdasarkan Tabel 2 dan Tabel 3, dapat dilihat bahwa ada korelasi antar variabel. Variabel-variabel yang memiliki korelasi linier tertinggi adalah variabel jumlah penduduk yang mempunyai keluhan kesehatan (X<sub>1</sub>) dengan jumlah penduduk yang mempunyai keluhan kesehatan dan berobat jalan (X<sub>2</sub>) sebesar 0,990. Kemudian juga variabel jumlah penduduk yang mempunyai keluhan kesehatan (X<sub>1</sub>) dengan jumlah penduduk yang pernah rawat inap (X<sub>3</sub>), variabel jumlah penduduk yang mempunyai keluhan kesehatan (X<sub>1</sub>) dan variabel jumlah penduduk yang menggunakan jaminan kesehatan untuk rawat inap (X<sub>4</sub>), variabel jumlah penduduk yang mempunyai keluhan kesehatan dan berobat jalan (X<sub>2</sub>) dengan jumlah penduduk yang pernah rawat inap (X<sub>3</sub>), (Jumlah penduduk yang mempunyai keluhan kesehatan dan berobat jalan (X<sub>2</sub>) dengan jumlah penduduk yang menggunakan jaminan kesehatan untuk rawat inap (X<sub>4</sub>), jumlah penduduk yang pernah rawat inap (X<sub>3</sub>) dengan jumlah penduduk yang menggunakan jaminan kesehatan untuk rawat inap (X<sub>4</sub>) yaitu masing-masing sebesar 0,950, 0,916, 0,972, 0,944, 0,991. Karena terdapat korelasi yang kuat antar variabel maka diperlukan analisis komponen utama. Sehingga korelasi antar variabel dapat dihilangkan agar asumsi dari analisis *cluster* terpenuhi.

### 3.2 Analisis Komponen Utama

Hasil dari Analisis Komponen Utama dapat dilihat pada Tabel 4 berikut:

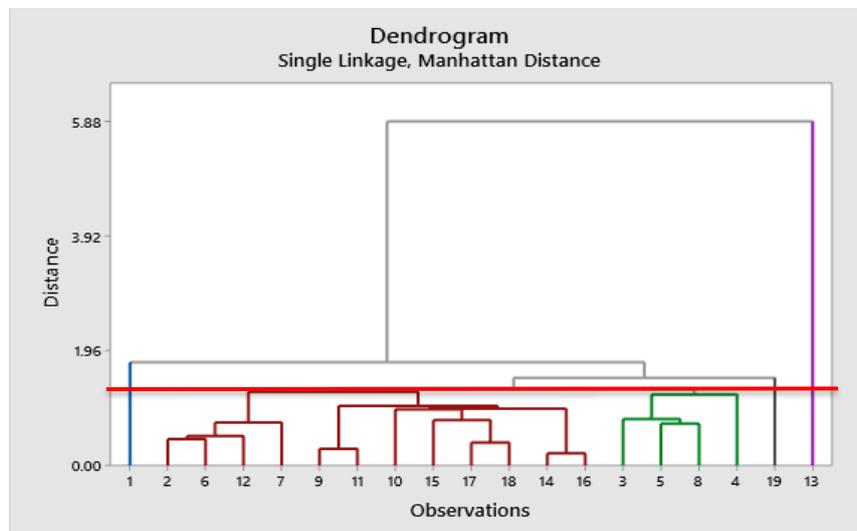
Tabel 4. Analisis Komponen Utama

Komponen Utama (KU)	Akar Ciri	Proporsi Keragaman	Keragaman Komulatif
KU1	3,7681	0,780	0,780
KU2	1,1163	0,197	0,977
KU3	0,1057	0,021	0,998
KU4	0,0064	0,001	0,999
KU5	0,0035	0,001	1,000

Salah satu cara dalam menentukan banyak komponen utama yang digunakan adalah dari melihat nilai eigen yang lebih dari 1 [4]. Pada Tabel 4 di atas diperoleh komponen utama yang memiliki nilai eigen yang lebih besar dari 1 adalah sebanyak 2 komponen utama yaitu komponen utama pertama dengan nilai eigen 3,7681 dan komponen kedua dengan nilai eigen 1,1163. Setelah analisis komponen utama dilakukan maka diperoleh skor komponen utama. Skor komponen utama tersebut menjadi peubah baru yang tidak saling berkorelasi. Sehingga dapat dilanjutkan ke analisis *cluster*.

### 3.3 Melakukan Analisis *Cluster Hierarki Metode Single Linkage*

Jumlah data yang digunakan pada penelitian ini berukuran kecil maka dilakukan analisis *cluster* dengan menggunakan metode hierarki [6], sehingga memudahkan dalam membentuk sebuah pohon tingkatan (hierarki) yang jelas antar objek dan belum diketahui sebelumnya banyaknya *cluster* yang akan dibentuk, serta menggunakan metode *single linkage* dan ukuran jarak manhattan. Hasil dari analisis *cluster* dapat dilihat dalam dendrogram pada Gambar 1.



Gambar 1. Dendrogram Hasil Pengklasteran Kabupaten/Kota di Sumatera Barat

### 3.4 Pemotongan Dendrogram

Pemotongan dendrogram pada Gambar 1 yang dilakukan berdasarkan subjektifitas terbagi atas lima *cluster*. *Cluster* pertama terdiri dari 1 kabupaten yaitu Kabupaten Kepulauan Mentawai. *Cluster* kedua terdiri dari 12 kabupaten/kota yaitu Kabupaten Pesisir Selatan, Kabupaten Padang Pariaman, Kabupaten Pasaman Barat, Kabupaten Agam, Kabupaten Pasaman, Kabupaten Dharmasraya, Kabupaten Solok Selatan, Kota Sawahlunto, Kota Bukittinggi, Kota Payakumbuh, Kota Solok dan Kota Padang Panjang. *Cluster* ketiga terdiri dari 4 kabupaten yaitu Kabupaten Solok, Kabupaten Tanah Datar, Kabupaten Lima Puluh Kota dan Kabupaten Sijunjung. *Cluster* keempat terdiri dari 1 kota yaitu Kota Pariaman. *Cluster* kelima terdiri dari 1 kota yaitu Kota Padang.

### 3.5 Interpretasi *Profiling Cluster*

Setelah mendapatkan jumlah anggota masing-masing *cluster* maka langkah selanjutnya adalah menginterpretasikan *profil* *cluster* yang bertujuan untuk tiap-tiap *cluster* menghasilkan karakteristik masing-masing. Dihitung rata-rata nilai objek yang ada dalam masing-masing *cluster* pada tiap variabel indikator kesehatan yang dapat dilihat pada Tabel 5.

Tabel 5. Nilai Rata-rata Setiap *Cluster*

Variabel	<i>Cluster</i> 1	<i>Cluster</i> 2	<i>Cluster</i> 3	<i>Cluster</i> 4	<i>Cluster</i> 5
X <sub>1</sub>	24.330	78.817,50333	<b>109.410,64</b>	29.320	<b>302.357</b>
X <sub>2</sub>	10.304	43.413,62689	<b>63.338,62184</b>	20.187	<b>188.882</b>
X <sub>3</sub>	1.141	12.426,93833	<b>16.816,19</b>	5.558	<b>59.836</b>
X <sub>4</sub>	830	9.173,506355	<b>12.518,674</b>	4.400	<b>50.927</b>
X <sub>5</sub>	3,88	4,703333333	<b>5,4025</b>	<b>6,15</b>	4,78

Berdasarkan Tabel 5 pada *cluster* 1 dapat dilihat bahwa semua variabelnya memiliki nilai dibawah rata-rata umum yang artinya kabupaten pada *cluster* ini yaitu Kabupaten Kepulauan Mentawai termasuk daerah dengan tingkat kesehatan yang tinggi. *Cluster* kedua semua variabelnya hampir mendekati rata-rata umum yang artinya pada *cluster* ini termasuk daerah dengan tingkat kesehatan tergolong sedang. *Cluster* ketiga semua variabelnya di atas rata-rata umum yang artinya pada *cluster* ini termasuk daerah dengan tingkat kesehatan tergolong rendah. *Cluster* keempat hampir semua variabelnya di bawah rata-rata umum. Artinya pada *cluster* ini termasuk daerah dengan tingkat kesehatan tergolong tinggi. *Cluster* kelima hampir semua variabelnya di atas rata-rata umum. Artinya pada *cluster* ini termasuk daerah dengan tingkat kesehatan tergolong rendah. Oleh karena itu, kabupaten atau kota pada *cluster* ketiga dan *cluster* kelima memerlukan perhatian yang lebih serius dari pihak-pihak pengambil kebijakan.

#### 4. Kesimpulan

*Cluster* pertama mencakup Kabupaten Kepulauan Mentawai yang termasuk daerah dengan tingkat kesehatan tinggi. *Cluster* kedua mencakup 12 kabupaten atau kota yaitu Kabupaten Pesisir Selatan, Kabupaten Padang Pariaman, Kabupaten Pasaman Barat, Kabupaten Agam, Kabupaten Pasaman, Kabupaten Dharmasraya, Kabupaten Solok Selatan, Kota Sawahlunto, Kota Bukittinggi, Kota Payakumbuh, Kota Solok dan Kota Padang Panjang termasuk daerah tingkat kesehatan sedang. *Cluster* ketiga terdiri dari 4 kabupaten yaitu Kabupaten Solok, Kabupaten Tanah Datar, Kabupaten Lima Puluh Kota dan Kabupaten Sijunjung termasuk daerah dengan tingkat kesehatan rendah. *Cluster* keempat terdiri dari 1 kota yaitu Kota Pariaman termasuk daerah tingkat kesehatan tinggi. *Cluster* kelima terdiri dari 1 kota yaitu Kota Padang termasuk daerah tingkat kesehatan rendah.

#### REFERENSI

- [1] Badan Pusat Statistik. (2021). *Profil Kesehatan Provinsi Sumatera Barat 2020*. Sumatera Barat: Badan Pusat Statistik Provinsi Sumatera Barat.
- [2] WHO. 2002. *Report on Strengthening National Health Information Systems: Workshop on the use of Geographical Information System Health Map*. Brazzaville: World Health Organization Regional Office for Africa.
- [3] Sitepu, R., Irmeilyana, I., & Gultom, B. (2011). Analisis *Cluster* terhadap Tingkat Pencemaran Udara pada Sektor Industri di Sumatera Selatan. *Jurnal Penelitian Sains*, 14(3).
- [4] Matjik, A. A. dan I Made S. (2011). *Sidik Peubah Ganda Dengan menggunakan SAS*. Bogor: IPB Press.
- [5] Mongi, C. E. (2015). Penggunaan Analisis *Two Step Clustering* untuk Data Campuran. *d'CARTESIAN*, 4(1), Hal 10-11.
- [6] Wirdiastuti, C., & Helma, H. (2019). Pengelompokan Sembilan Kabupaten/Kota di Provinsi Sumatera Barat Berdasarkan Tingkat Kriminalitas dengan Menggunakan Analisis Gerombol.



---

*UNP: Journal of Mathematics*, 4(2).