

Original Article

e-ISSN: 2774-2016 - <https://journal.itera.ac.id/index.php/indojam/>

p-ISSN: 2774-2067

Received 5th March 2024

Accepted 8th May 2024

Published 30th May 2024

Open Access

DOI:

<https://doi.org/10.35472/indoja.m.v4i1.1772>

Analisis Jumlah Kecelakaan Lalu Lintas, Korban, dan Kerugian di Provinsi Lampung Tahun 2021 dengan Metode K-Means Cluster

Muklas Riva'i^a, Aulia Khairani Hutabarat^a, Doni Hardian Halawa^a, Ilham Satrio^a, Muhammad Syauki Tarmizi^a, Nadia Frasiska^a, Niken Pratiwi^a, Zalikha Alma^a

^a Program Studi Aktuaria, Fakultas Sains, Institut Teknologi Sumatera

* Koresponden E-mail: muklas.rivai@at.itera.ac.id

Abstract: Lampung Province has a relatively high rate of traffic accidents. These accidents are caused by the increasing number of vehicles and inadequate road infrastructure. Apart from that, public awareness regarding traffic regulations is relatively low. The aim of this research is to group traffic accident data in Lampung Province in 2021. The method used is the Cluster Analysis method in Geographic Information Systems (GIS). The results of the analysis show that the clusters formed in the number of traffic accidents, victims and losses in Lampung Province in 2021 are 2 clusters. Cluster 1 consists of West Lampung, Pesawaran, Mesuji, West Tulang Bawang, Bandar Lampung, Metro and Cluster 2 consists of Tanggamus, South Lampung, East Lampung, Central Lampung, North Lampung, Way Kanan, Tulang Bawang, Pringsewu. Then the characteristics of Cluster 1 are that the road conditions are relatively good and there is road condition maintenance that covers all existing roads. Meanwhile, Cluster 2 has a high road length and is crossed by the Sumatra route.

Keywords: *K-means, Accidents, Lampung Province*

Abstrak: Provinsi Lampung memiliki tingkat kecelakaan lalu lintas yang relatif tinggi. Kecelakaan tersebut disebabkan oleh meningkatnya jumlah kendaraan yang semakin tinggi dan infrastruktur jalan raya yang belum memadai. Selain itu juga, kesadaran masyarakat terkait peraturan lalu lintas yang relatif rendah. Tujuan Penelitian ini adalah untuk mengelompokkan data kecelakaan lalu lintas di Provinsi Lampung pada tahun 2021. Metode yang digunakan yaitu metode Cluster Analysis dalam Sistem Informasi Geografis (SIG). Hasil analisis menunjukkan bahwa cluster yang terbentuk pada Jumlah Kecelakaan Lalu Lintas, Korban, dan kerugian di Provinsi Lampung Tahun 2021 adalah 2 cluster. Cluster 1 terdiri dari Lampung Barat, Pesawaran, Mesuji, Tulang Bawang Barat, Bandar Lampung, Metro dan Cluster 2 terdiri dari Tanggamus, Lampung Selatan, Lampung Timur, Lampung Tengah, Lampung Utara, Way Kanan, Tulang Bawang, Pringsewu. Kemudian karakteristik Cluster 1 kondisi jalan relatif baik dan terdapat pemeliharaan kondisi jalan yang menjangkau semua jalan yang ada. Sedangkan Cluster 2 memiliki Panjang jalan yang tinggi serta dilewati oleh jalur lintas Sumatera.

Kata Kunci: *K-means, Kecelakaan, Provinsi Lampung*

Pendahuluan

Saat ini, Provinsi Lampung menghadapi kondisi jalan raya yang cukup memprihatinkan. Meningkatnya jumlah kendaraan serta kurangnya sarana dan prasarana yang memadai mengakibatkan angka

kecelakan di Provinsi Lampung meningkat. Ada banyak faktor yang menyebabkan kecelakaan di jalan raya, di antaranya meningkatnya jumlah penduduk dan kurangnya kesadaran dalam menaati peraturan lalu lintas [1].

Original Article

Kenyataan yang diterima oleh masyarakat saat ini adalah pemerintah yang kurang memberikan perhatian serius terhadap kerusakan infrastruktur. Hal ini dibuktikan dengan angka data kecelakaan lalu lintas yang kian meningkat tiap tahun. Namun, kecelakaan lalu lintas yang diakibatkan oleh kerusakan jalan tersebut bukan merupakan satu-satunya faktor utama. Akan tetapi, faktor utama dari kecelakaan lalu lintas tersebut adalah kelalaian dari pengemudi [2].

Kecelakaan lalu lintas berdasarkan Pasal 1 ayat 24 UU No 22 Tahun 2009 Tentang Lalu Lintas dan Angkutan Jalan (UU LLAJ). Kecelakaan lalu lintas adalah suatu peristiwa di jalan yang tidak diduga dan tidak disengaja melibatkan kendaraan dengan atau tanpa pengguna jalan lain yang mengakibatkan korban manusia dan/atau kerugian harta benda [3]. Beberapa faktor penyebab kecelakaan lalu lintas, yaitu faktor pemakai jalan, faktor jalan, faktor cuaca, dan faktor kendaraan [4].

Sistem Informasi Geografis (SIG) adalah sistem informasi pemetaan berbasis komputer yang digunakan untuk memasukkan, mengolah, menganalisis, dan menghasilkan data yang bereferensi geografis sehingga data tersebut dapat digunakan untuk mendukung pengambilan keputusan yang tepat berdasarkan kondisi geografis tertentu. Salah satu metode dalam SIG yang digunakan dalam menganalisis serta mengelompokkan suatu data adalah *Cluster Analysis*.

Pada penelitian ini, metode *Cluster Analysis* digunakan karena mampu mengelompokkan data observasi dalam jumlah besar dan variabel yang relatif banyak sehingga data yang direduksi dapat dikelompokkan dengan mudah serta dengan metode ini akan mendapatkan nilai akurasi lebih tinggi dibanding analisis *cluster* lainnya [5]. Salah satu dari teknik clustering untuk data numerik adalah K-Means yang merupakan metode clustering yang membagi data ke dalam k cluster untuk meminimalkan varians

dalam cluster. K-Means mudah untuk diimplementasikan dan juga memiliki kompleksitas waktu dan ruang yang kecil sehingga hasil yang didapat cukup baik dan memuaskan [6].

Analisis cluster terdiri dari metode hierarki dan non hierarki. Metode hierarki digunakan jika belum ada informasi jumlah cluster yang akan dipilih. [7] K-Means adalah salah satu metode analisis cluster non hierarki. Pengklasteran K-Means sangat sesuai jika digunakan pada data dengan ukuran yang besar dikarenakan mempunyai perhitungan yang cepat [8]. Namun, metode K-Means terbatas pada data numerik. Sehingga dibutuhkan metode untuk mengolah data kategorik. Salah satu metode non hierarki untuk data kategorik yang sering digunakan adalah K-Modes. K-Means adalah salah satu metode yang digunakan dalam clustering dengan tujuan untuk menganalisis suatu data dengan melakukan proses pemodelan tanpa supervise (unsupervised) dan metode tersebut juga dapat mengelompokkan data dengan system partisi [9].

Dengan menggunakan metode K-Means, variasi antar data yang terdapat dalam suatu cluster dapat diminimumkan dan variasi dengan data yang ada di cluster lainnya dapat dimaksimalkan. Selain itu, K-Means mempunyai kemampuan mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang relative cepat dan efisien.

Dalam hal ini, pengambilan data cluster menjadi penting sebagai langkah awal untuk memahami pola-pola kecelakaan lalu lintas di Provinsi Lampung. Pengumpulan data cluster pada tingkat kecelakaan dapat memberikan wawasan yang mendalam tentang faktor-faktor utama yang menyebabkan kecelakaan, pola kejadian, serta karakteristik khusus dari lokasi-lokasi tertentu di wilayah tersebut.

Metode

Data

Pada penelitian ini digunakan data sekunder. Data sekunder merupakan sumber data yang sudah ada sebelumnya. Data yang digunakan adalah data pada tahun 2021 [10]. Data tersebut diperoleh dari BPS (Badan Pusat Statistika) melalui arsip yang telah dibuat dan dipublikasikan oleh lembaga-lembaga terkait, yang kemudian akan diolah melalui RStudio. Arsip data yang diperoleh dan akan digunakan pada penelitian ini merupakan data dari Provinsi Lampung. Dalam hal ini terdapat 14 Kabupaten/Kota antara lain, Lampung Barat, Tanggamus, Lampung Selatan, Lampung Timur, Lampung Tengah, Lampung Utara, Way Kanan, Tulang Bawang, Pesawaran, Pringsewu, Mesuji, Tulang Bawang Barat, Bandar Lampung, dan Metro. Serta variabel yang digunakan dalam penelitian ini ada 5 variabel yakni jumlah kecelakaan (X_1), jumlah korban meninggal dunia (X_2), jumlah korban yang mengalami luka berat (X_3), jumlah korban yang mengalami luka ringan (X_4), dan jumlah kerugiannya (X_5). Variabel-variabel tersebut akan digunakan sebagai faktor-faktor untuk analisis *clustering* terhadap Provinsi Lampung yang dalam hal ini terbagi ke 14 Kabupaten/Kota.

Data penelitian pada tahun 2021 di Provinsi Lampung yang diperoleh dari BPS (Badan Pusat Statistik) ditunjukkan pada Tabel berikut:

TABEL 1 Data Penelitian Tahun 2021

Polres/ Polresta Kabupaten	X_1	X_2	X_3	X_4	X_5 (dlm juta)
Lampung Barat	98	41	13	70	158,1
Tanggamus	99	35	36	116	347,6
Lampung Selatan	279	101	236	233	2772,95
Lampung Timur	193	103	69	157	762,8

Lampung Tengah	214	101	163	109	1751,4
Lampung Utara	176	39	82	210	939,1515
Way Kanan	88	27	54	72	375,4
Tulang Bawang	87	35	45	84	550,6
Pesawaran	74	35	9	102	369,75
Pringsewu	91	25	53	91	154,25
Mesuji	31	29	4	31	295,3
Tulang Bawang Barat	44	25	20	22	444,5
Bandar Lampung	175	63	8	197	453,5
Metro	89	9	1	123	107

Prosedur penelitian

Berikut adalah langkah-langkah Prosedur penelitian antara lain [11]:

1. Mengumpulkan data dari BPS Lampung dalam Data 2022
2. Melakukan eksplorasi data dalam bentuk grafik maupun sari numerik untuk menampilkan informasi analisa awal
3. Melakukan analisis algoritma clustering K- means dengan langkah sebagai berikut.
 - a) Menentukan nilai k sebagai jumlah kluster yang akan di bentuk. Nilai k awal yang dibentuk adalah k=2.
 - b) Membangkitkan nilai random untuk pusat *cluster* awal (*centroid*) sebanyak k-*cluster*
 - c) Membangkitkan jarak setiap data input terhadap masing-masing *centroid* menggunakan rumus jarak *Euclidian* (*Euclidian Distance*) hingga ditemukan jarak yang paling dekat dari setiap data dengan *centroid*. Berikut adalah persamaan *Euclidian Distance*:

$$d(x_i, \mu_i) = \sqrt{(x_i - \mu_i)^2}$$

Original Article

Dengan $d(xi, \mu_i)$ adalah jarak antara *cluster* x dengan pusat *cluster* μ pada kata ke i , xi adalah bobot kata ke I , pada *cluster* yang ingin dicari jaraknya, μ_i bobot kata ke i pada pusat *cluster*.

- d) Mengklasifikasi setiap data berdasarkan kedetannya dengan *centroid* (jarak terkecil).
 - e) Mengupdate nilai *centroid*. Nilai *centroid* baru diperoleh dari rata-rata *cluster* yang bersangkutan dengan menggunakan rumus:
Dimana:

$$C_k = \frac{1}{n_k} \sum d_i$$

n_k = jumlah data dalam *cluster*
 d_i = jumlah dari nilai jarak yang masuk dalam masing-masing *cluster*
 - f) Melakukan perulangan dari Langkah 2 hingga 5 sampai tiap *cluster* tidak ada yang berubah.
 - g) Jika Langkah f) telah terpenuhi, maka nilai rata-rata pusat *cluster* (μ_i) dan akan digunakan sebagai parameter untuk menentukan klasifikasi data.
4. Membuat grafik plot clustering
 5. Melakukan Interpretasi dan analisis terhadap hasil penelitian
 6. Menarik kesimpulan dan saran terhadap hasil penelitian

Hasil dan Diskusi

Statistika Deskriptif

Tingkat kecelakaan di Provinsi Lampung pada tahun 2021 relatif tinggi. Tingkat kecelakaan terendah di tahun 2021 yaitu kabupaten Mesuji dengan 31 kasus. Kemudian tingkat kecelakaan tertinggi di tahun 2021 yaitu kabupaten Lampung Selatan dengan 279 kasus. Hal ini disebabkan karena Kabupaten Lampung Selatan menjadi gerbang pertama bagi kendaraan-

kendaraan dari pulau jawa yang akan menuju pulau Sumatera. Berikut adalah statistik deskriptif dari data yang digunakan:

TABEL 2 Tabel Analisis Deskriptif

	X ₁	X ₂	X ₃	X ₄	X ₅ (d/m juta)
Min	31,00	9,00	1,00	22,0	107
Mean	124,14	47,71	56,64	115,5	677,3
Max	279,00	103,0	236,00	233,0	277,3
Std. Dev	71,26	31,48	67,14	63,7	736,4

Uji Multikolinearitas

Sebelum melakukan analisis *clustering* terlebih dahulu dilakukan uji multikolinearitas untuk melihat apakah terdapat hubungan korelasi antar variabel dengan menggunakan nilai VIF. Berikut merupakan uji hipotesis dari uji multikolinearitas:

Hipotesis:

H_0 = Tidak terjadi multikolinearitas

H_1 = Terjadi multikolinearitas

TABEL 3 Uji Multikolinearitas

	Collinearity Tolerance	Statistics VIF
X ₁ (constant)		
X ₂	0,384	2,605
X ₃	0,077	13,008
X ₄	0,629	1,591
X ₅	0,065	15,286

Berdasarkan tabel diperoleh VIF dan nilai *tolerance* yang diberikan antar variabel persentase jumlah kecelakaan, jumlah kematian, dan luka ringan memiliki nilai *tolerance* lebih dari 0,1 dan nilai VIF < 10. Sedangkan variabel luka berat, dan total kerugian memiliki nilai *tolerance* kurang dari 0,1 dan nilai VIF > 10. Karena masih ada nilai *tolerance* yang kurang dari 0,1 dan nilai VIF > 10, maka diperoleh keputusan tidak tolak H_0 [12]. Sehingga, dapat disimpulkan bahwa masih ada data yang terjadi multikolinearitas. Karena

data terjadi multikolinearitas, maka dilakukan transformasi data ke dalam bentuk logaritma natural. Berikut ini merupakan tabel hasil uji multikolinearitas dengan transformasi data:

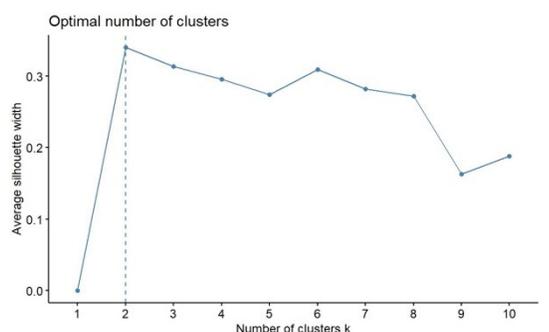
TABEL 3 Hasil Uji Multikolinearitas

	Collinearity Tolerance	Statistics VIF
X_1 (constant)		
LN_ X_2	0,324	3,085
LN_ X_3	0,399	2,506
LN_ X_4	0,809	1,236
LN_ X_5	0,276	3,618

Berdasarkan tabel di atas, didapatkan bahwa semua variabel memiliki nilai *tolerance* lebih besar dari 0,1 serta nilai VIF nya < 10 . Jadi dapat disimpulkan bahwa semua variabel dengan data bentuk logaritma natural tidak terjadi multikolinearitas.

Hasil perhitungan K-Means Clustering

Pada analisis *cluster* untuk mengetahui bagaimana pengelompokan Kabupaten/Kota di Provinsi Lampung berdasarkan tingkat kecelakaan pada tahun 2021 perlu dilakukan penentuan jumlah *cluster* dengan pendekatan *Average Silhouette Width* untuk mendapatkan *cluster* optimal. Berikut hasil *cluster* optimal yang didapatkan:



Gambar 1 cluster optimal

Berdasarkan gambar hasil pendekatan Silhouette didapatkan nilai k optimal pada saat nilai $k = 2$,

ditandai dengan melihat garis Silhouette pada $k = 2$. Sesuai pada tujuan penelitian ini akan digunakan $k = 2$ yaitu untuk mengelompokkan kabupaten di provinsi Lampung berdasarkan Jumlah kematian, meninggal dunia, luka berat, luka ringan, dan jumlah kerugian.

Hasil dan Analisis Cluster

Pada perhitungan analisis *cluster* menggunakan pendekatan *k-means* dilakukan proses iterasi hingga jumlah anggota *cluster* iterasi maksimal (jumlah *cluster* setiap iterasi sama). Proses iterasi pertama dilakukan dengan menggunakan rumus korelasi antara tiga objek agar mengetahui jarak tiap data ke pusat *centroid*. Nilai *Euclidian Distance* yang didapatkan setiap kabupaten dengan *centroid* adalah sebagai berikut:

TABEL 4 Hasil Perhitungan Analisis Cluster

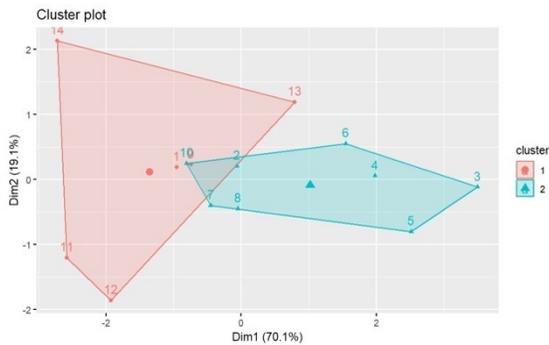
Kode Kabupaten	Polres/Polresta Kabupaten	Cluster	Distance
1	Lampung Barat	1	0.984
2	Tanggamus	2	1.094
3	Lampung Selatan	2	2.181
4	Lampung Timur	2	0.860
5	Lampung Tengah	2	1.502
6	Lampung Utara	2	0.720
7	Way Kanan	2	1.137
8	Tulang Bawang	2	0.883
9	Pesawaran	1	0.622
10	Pringsewu	2	1.728
11	Mesuji	1	1.283
12	Tulang Bawang Barat	1	1.773
13	Bandar Lampung	1	2.463
14	Metro	1	1.659

Original Article

Berdasarkan hasil *clustering* dengan menggunakan software SPSS dengan $k = 2$, didapatkan jumlah anggota *cluster* 1 memiliki jumlah data sebanyak 6 Kabupaten (Lampung Barat, Pesawaran, Mesuji, Tulang Bawang Barat, Bandar Lampung, Metro), dan *cluster* 2 memiliki jumlah data sebesar 8 Kabupaten (Tanggamus, Lampung Selatan, Lampung Timur, Lampung Tengah, Lampung Utara, Way Kanan, Tulang Bawang, Pringsewu).

Cluster Plot

Setelah mendapatkan hasil *cluster* dilakukan visualisasi hasil clusterisasi yang digunakan untuk melihat persebaran *cluster*. Berdasarkan gambar, diperoleh bahwa plot yang berwarna merah menunjukkan provinsi yang termasuk ke dalam *cluster* 1, plot berwarna biru menunjukkan provinsi yang termasuk ke dalam *cluster* 2.



Gambar 2 Cluster plot

Setelah dibentuk dalam bentuk plot dilakukan pengelompokan data *cluster* dalam bentuk tabel untuk melihat provinsi mana saja yang terbentuk kedalam masing-masing *cluster*. Berikut adalah hasil analisis menggunakan software SPSS.

TABEL 5 Kabupaten/Kota Tiap cluster

Cluster	Kabupaten	Jumlah
1	Lampung Barat, Pesawaran, Mesuji, Tulang Bawang Barat, Bandar Lampung, Metro	6
2	Tanggamus, Lampung Selatan, Lampung Timur, Lampung Tengah, Lampung Utara, Way Kanan, Tulang Bawang, Pringsewu.	8

• Karakteristik Cluster 1

Berdasarkan tabel diatas menunjukkan bahwa *cluster* satu pada variabel jumlah kecelakaan, meninggal dunia, luka berat, luka ringan, dan jumlah kerugian memiliki tingkat yang rendah dibanding *cluster* kedua. Selain itu juga, kabupaten yang berada di *cluster* pertama memiliki kondisi jalan relatif lebih baik, kemudian pemeliharaan sarana prasarana jaringan jalan cukup memadai oleh pemerintah daerahnya.

• Karakteristik Cluster 2

Berdasarkan tabel diatas menunjukkan bahwa *cluster* dua pada variabel jumlah kecelakaan, meninggal dunia, luka berat, luka ringan, dan jumlah kerugian memiliki tingkat yang tinggi dibanding *cluster* satu. Selain itu juga, Panjang jaringan jalan yang dibebankan kabupaten tersebut relatif tinggi sehingga anggaran pemerintah daerah tidak mencukupi untuk Pembangunan dan perawatan jaringan jalan. Kemudian Sebagian besar kabupaten pada cluster kedua dilewati jalur lintas Sumatera sehingga banyak kendaraan roda empat yang melewati jalur tersebut yang merupakan kendaraan antar provinsi yang kemungkinan terjadinya kendala pada mobil atau supir relative tinggi.

Kesimpulan

Dari hasil analisis *cluster* terhadap data kecelakaan lalu lintas di Provinsi Lampung pada tahun 2021, dapat diambil kesimpulan bahwa: terdapat 2 cluster yang terbagi sebagai berikut. *Cluster* 1 terdiri dari Lampung Barat, Pesawaran, Mesuji, Tulang Bawang Barat, Bandar Lampung, Metro dan *Cluster* 2 terdiri dari Tanggamus, Lampung Selatan, Lampung Timur, Lampung Tengah, Lampung Utara, Way Kanan, Tulang Bawang, Pringsewu. Kemudian karakteristik *Cluster* 1 kondisi jalan relatif baik dan terdapat pemeliharaan kondisi jalan yang menjangkau semua jalan yang ada. Sedangkan *Cluster* 2 memiliki Panjang

jalan yang tinggi serta dilewati oleh jalur lintas Sumatera.

Referensi

- [1] M. Ir. Lilis Widodo, "Penerapan Metode Indeks Bahaya Kecelakaan Untuk Analisis Kasus Lalu Lintas Di Bandar Lampung," *Jurnal Teknik Sipil Ubl*, Vol. 7, P. 873, 2016.
- [2] R. L. Jannah, H. Yermadona, and S. Dewi, "Analisis Kerusakan Perkerasan Jalan Dengan Metoda Bina Marga Dan Pavement Condition Index (PCI)(Studi kasus: Jl. Lintas Sumatera)," *Ensiklopedia Res. Community Serv. Rev.*, vol. 1, no. 2, pp. 114–122, 2022.
- [3] P. Indonesia, "Undang-Undang Republik Indonesia Nomor 22 Tahun 2009," 2009.
- [4] A. L. N. M. A. Nanda Dewi Arumsari, "Pemodelan Daerah Rawan Kecelakaan Dengan Analysis Cluster," p. 10, 2016.
- [5] M. W. Talakua, Z. A. Leleury and A. W. Talluta, "Analisis Cluster Dengan Menggunakan Metode K-Means Untuk Pengelompokan Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia Tahun 2014," *Jurnal Ilmu Matematika Dan Terapan*, vol. 11, pp. 119-128, 2017.
- [6] C. S. D. Sembiring, L. Hanum, dan S. P. Tamba, "Penerapan Data Mining Menggunakan Algoritma K-Means untuk Menentukan Judul Skripsi dan Jurnal Penelitian (Studi Kasus FTIK UNPRI)," *Jurnal Sistem Informasi dan Ilmu Komputer Prima*, vol. 5, no. 2, hlm. 80–85, 2022.
- [7] J. Han and J. P. Michelinw Kamber, *Data Mining Concepts and Techniques*, ELSEVIER, 2012.
- [8] A. Y. Ananta, "Seleksi Notifikasi Serangan Berbasis IDS Snort Menggunakan Metode K-Means," *Jurnal Unikama*, vol. 3, p. 31, 2017.
- [9] C. C. Aggarwal and C. K. Reddy, *Data Clustering Algorithms and Applications*, CRC Press, 2014.
- [10] N. Martono, *Metode Penelitian Kuantitatif*, Jakarta: Rajagrafindo Persada, 2014.
- [11] S. W. S. E. S. Lynda Rahmawati, "Analisa Clustering Menggunakan Metode K-Means Dan Hierarchical Clustering," 10.
- [12] M. Sriningsih, D. Hatidja and J. D. Prang, "Penanganan Multikolinearitas Dengan Menggunakan Analisis Regresi Komponen Utama Pada Kasus Impor Beras Diprovinsi Sulut," *Jurnal Ilmiah Sains*, vol. 18, pp. 18-24, 2018.