**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 9601 – 9610 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

# PENGELOMPOKAN KABUPATEN/KOTA BERDASARKAN PRODUKSI TANAMAN BUAH TAHUNAN MENGGUNAKAN METODE K-MEANS DI PROVINSI SUMATERA UTARA

**Junita Sastri [1], Helmi Fauzi Siregar [2]**
**Universitas Asahan, Asahan**
e-mail: [1] junitasastri19@gmail.com, [2] fauzi.helmi.hf@gmail.com

***Abstract*** *: Differences in annual fruit production across regencies and cities in North Sumatra require systematic analysis to identify areas with similar production characteristics. This study aims to group regencies and cities in North Sumatra Province based on avocado and starfruit production using the K-Means klastering method. The study applies a quantitative approach using secondary data published by Statistics Indonesia (BPS) of North Sumatra Province for 2022–2025. The dataset consists of 132 observations representing 33 regencies/cities over four years, with two variables: avocado production and starfruit production. The K-Means process includes determining three klasters, initializing centroids, calculating Euclidean distances, assigning observations to the nearest centroid, and updating centroids until convergence. The final klastering converged at the sixth iteration. The results produced three klasters: high production (12 observations), medium production (20 observations), and low production (100 observations). The largest group was the low-production klaster, indicating that most observations had comparatively lower production characteristics for the two commodities. A web-based application developed using PHP and MySQL supports data management, klastering, result presentation, and report generation. The results can provide preliminary information for understanding the distribution of avocado and starfruit production across regions in North Sumatra.*

***Keywords:*** *K-Means, klastering, annual fruit production, avocado, starfruit, North Sumatra.*

**Abstrak** : Perbedaan produksi tanaman buah tahunan pada setiap kabupaten/kota di Provinsi Sumatera Utara perlu dianalisis secara sistematis untuk mengetahui wilayah yang memiliki karakteristik produksi serupa. Penelitian ini bertujuan mengelompokkan kabupaten/kota berdasarkan produksi buah alpukat dan belimbing menggunakan metode K-Means. Penelitian menggunakan pendekatan kuantitatif dengan data sekunder yang bersumber dari publikasi Badan Pusat Statistik (BPS) Provinsi Sumatera Utara periode 2022–2025. Data terdiri atas 132 pengamatan dari 33 kabupaten/kota selama empat tahun, dengan dua variabel yaitu produksi alpukat dan produksi belimbing. Tahapan K-Means meliputi penentuan tiga klaster, inisialisasi centroid, perhitungan jarak Euclidean, penentuan keanggotaan berdasarkan jarak terdekat, dan pembaruan centroid hingga konvergen. Hasil perhitungan mencapai kondisi konvergen pada iterasi keenam dan menghasilkan tiga klaster, yaitu produksi tinggi sebanyak 12 data, produksi sedang sebanyak 20 data, dan produksi rendah sebanyak 100 data. Klaster produksi rendah memiliki jumlah anggota terbesar, yang menunjukkan bahwa sebagian besar pengamatan memiliki karakteristik produksi kedua komoditas yang relatif rendah dibandingkan kelompok lainnya. Metode ini diimplementasikan dalam aplikasi berbasis web menggunakan PHP dan MySQL untuk membantu pengelolaan data, proses pengelompokan, penyajian hasil, dan pembuatan laporan. Hasil penelitian diharapkan dapat menjadi informasi awal untuk memahami persebaran produksi alpukat dan belimbing di Provinsi Sumatera Utara.

**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 9601 – 9610 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

**Kata kunci:** K-Means, klastering, produksi buah tahunan, alpukat, belimbing, Sumatera Utara.

## PENDAHULUAN

Sektor pertanian memiliki peran penting dalam mendukung perekonomian dan memenuhi kebutuhan masyarakat. Salah satu subsektor pertanian adalah hortikultura, termasuk tanaman buah tahunan yang memiliki nilai ekonomi serta peluang pengembangan di berbagai wilayah. Provinsi Sumatera Utara memiliki keragaman produksi buah antarkabupaten/kota. Perbedaan jumlah produksi tersebut menyebabkan data statistik perlu diolah agar pola dan kemiripan karakteristik antarwilayah dapat diketahui secara lebih mudah (Annisa Wulandari, 2025).

Data produksi yang disajikan dalam bentuk angka belum secara langsung memperlihatkan kelompok wilayah dengan karakteristik produksi yang serupa. Pengelompokan data dapat membantu merangkum informasi tersebut dan memberikan gambaran awal mengenai persebaran produksi buah. Salah satu teknik yang dapat digunakan adalah data mining melalui klastering, yaitu proses pengelompokan objek berdasarkan kemiripan karakteristiknya (Nilam Kumala, 2023).

K-Means merupakan salah satu algoritma klastering yang membagi data ke dalam sejumlah klaster berdasarkan jarak setiap data terhadap pusat klaster atau centroid. Setiap data ditempatkan pada klaster dengan jarak terdekat, kemudian centroid diperbarui berdasarkan anggota klaster. Proses tersebut dilakukan secara berulang hingga keanggotaan klaster tidak berubah atau mencapai kondisi konvergen.

Penelitian terdahulu oleh (Harun Mukhtar, 2024), menerapkan K-Means untuk menganalisis kesuburan pertanian melalui data irigasi. Sementara itu, (Risbuwono Heru Cokro, 2025), menerapkan K-Means untuk mengelompokkan provinsi penghasil padi berdasarkan karakteristik produksi. Penelitian-penelitian tersebut menunjukkan penggunaan klastering untuk menyederhanakan data pertanian ke dalam kelompok karakteristik tertentu. Penelitian ini berfokus pada pengelompokan kabupaten/kota di Provinsi Sumatera Utara berdasarkan produksi alpukat dan belimbing selama periode 2022–2025.

Tujuan penelitian ini adalah menerapkan metode K-Means untuk mengelompokkan data produksi alpukat dan belimbing ke dalam tiga kategori, yaitu tinggi, sedang, dan rendah, serta mengimplementasikan proses pengelompokan tersebut pada aplikasi berbasis web. Hasilnya diharapkan dapat membantu penyajian informasi produksi buah secara lebih terstruktur. Hasil klaster merupakan informasi pendukung dan tidak dimaksudkan untuk menjelaskan seluruh faktor yang memengaruhi produksi pertanian.

## METODE

Metode penelitian merupakan tahapan sistematis yang digunakan untuk menyelesaikan permasalahan penelitian agar proses penelitian berjalan secara terarah dan sesuai dengan tujuan yang telah ditetapkan.

Penelitian ini menggunakan metode K-Means untuk mengelompokkan kabupaten/kota berdasarkan produksi tanaman buah tahunan, yaitu alpukat dan belimbing, di Provinsi Sumatera Utara.

Tahapan penelitian yang dilakukan meliputi identifikasi masalah, pengumpulan data, pra-proses data, penerapan metode K-Means, perancangan sistem, pembuatan aplikasi web, pengujian dan evaluasi, serta penarikan kesimpulan. Kerangka kerja penelitian ditunjukkan pada Gambar 1.

**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 9601 – 9610 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR




**Gambar 1 Kerangka Kerja Penelitian**

Adapun uraian tahapan kerangka kerja penelitian adalah sebagai berikut. (1) Identifikasi masalah dilakukan untuk mengetahui permasalahan pengelompokan kabupaten/kota berdasarkan produksi alpukat dan belimbing di Provinsi Sumatera Utara. (2) Pengumpulan data dilakukan dengan memperoleh data sekunder dari publikasi BPS Provinsi Sumatera Utara tahun 2022–2025. (3) Pra-proses data dilakukan dengan mempersiapkan dan memeriksa data produksi sebelum diolah. (4) Penerapan metode K-Means dilakukan untuk mengelompokkan data menjadi tiga klaster berdasarkan kemiripan produksi menggunakan perhitungan Euclidean Distance. (5) Perancangan sistem dilakukan menggunakan pendekatan UML. (6) Pembuatan aplikasi web dilakukan untuk mengelola data, memproses pengelompokan, dan menampilkan Hasil klasterisasi. (7) Pengujian dan evaluasi dilakukan menggunakan metode Black-Box Testing untuk memeriksa fungsi aplikasi. (8) Penarikan kesimpulan dilakukan berdasarkan Hasil pengelompokan dan pengujian sistem.

**Pengumpulan Data**

Penelitian ini menggunakan pendekatan kuantitatif dan data sekunder yang diperoleh dari Badan Pusat Statistik Provinsi Sumatera Utara melalui publikasi statistik produksi tanaman buah tahunan. Data mencakup 33 kabupaten/kota pada periode 2022–2025. Dengan empat tahun pengamatan untuk setiap wilayah, jumlah keseluruhan data adalah 132 pengamatan. Variabel yang digunakan terdiri atas produksi alpukat sebagai X1 dan produksi belimbing sebagai X2, dengan satuan mengikuti publikasi sumber data.

**Tabel 1. Variabel Input Penelitian**

| Kode | Variabel | Jenis Data | Sumber |
|---|---|---|---|
| X1 | Produksi alpukat | Numerik | BPS Provinsi Sumatera Utara |
| X2 | Produksi belimbing | Numerik | BPS Provinsi Sumatera Utara |

**Penerapan Metode K-Means**

Metode K-Means digunakan untuk mengelompokkan data berdasarkan kemiripan nilai variabel produksi alpukat dan belimbing. Jumlah klaster yang digunakan adalah tiga, yaitu K1, K2, dan K3, yang selanjutnya diberi label tinggi, sedang, dan rendah sesuai dengan Hasil pengelompokan penelitian.

Tahapan penerapan K-Means meliputi penentuan jumlah klaster, penentuan centroid awal, perhitungan jarak Euclidean, penempatan setiap data pada klaster dengan jarak terdekat, pembaruan centroid, dan pengulangan proses sampai tidak terjadi perubahan keanggotaan klaster. Setelah jarak dihitung, setiap data ditempatkan ke klaster dengan nilai jarak terkecil. Centroid baru kemudian dihitung berdasarkan rata-rata nilai setiap variabel

dari anggota klaster. Proses tersebut diulangi sampai keanggotaan klaster tidak berubah atau proses mencapai kondisi konvergen.

## HASIL DAN PEMBAHASAN

### Analisis Data

Data penelitian berupa data produksi alpukat dan belimbing dari 33 kabupaten/kota di Provinsi Sumatera Utara selama periode 2022–2025. Setiap kombinasi kabupaten/kota dan tahun diperlakukan sebagai satu pengamatan, sehingga diperoleh 132 pengamatan. Penelitian ini menggunakan dua variabel, yaitu produksi alpukat sebagai X1 dan produksi belimbing sebagai X2. Kedua variabel tersebut digunakan sebagai dasar pengelompokan menggunakan metode K-Means untuk mengidentifikasi kelompok pengamatan yang memiliki karakteristik produksi serupa.

Kedua variabel tersebut digunakan sebagai dasar pengelompokan untuk melihat kemiripan pola produksi. Perbedaan nilai produksi antarwilayah dan antarperiode menjadi dasar bagi metode untuk menempatkan setiap catatan ke dalam klaster yang sesuai. Tabel 2.

**Tabel 2. Data Penelitian**

| No | Kabupaten/Kota | Tahun | Produksi Alpukat (kuintal) (Kw) | Produksi Belimbing (kuintal) (Kw) |
|---|---|---|---|---|
| 1 | Nias | 2022 | 122 | 48 |
| 2 | Mandailing Natal | 2022 | 11248 | 23993 |
| 3 | Tapanuli Selatan | 2022 | 120169 | 1349 |
| 4 | Tapanuli Tengah | 2022 | 62 | 8 |
| 5 | Tapanuli Utara | 2022 | 86650 | 2807 |
| 6 | Toba Samosir | 2022 | 13302 | 9 |
| 7 | Labuhan Batu | 2022 | 6 | 90 |
| 8 | Asahan | 2022 | 5675 | 2836 |
| 9 | Simalungun | 2022 | 13069 | 264 |
| 10 | Dairi | 2022 | 64185 | 0 |
| .. | ……. | ……. | …. | .. |
| … | …….. | …….. | ……. | … |
| 132 | Kota Gunungsitoli | 2025 | 1305 | 666 |

### Penerapan Metode K-Means

Penerapan metode K-Means dilakukan untuk mengelompokkan data produksi alpukat dan belimbing menjadi tiga klaster. Jumlah klaster ditetapkan sebanyak (K=3), yaitu klaster produksi tinggi, sedang, dan rendah. Proses pengelompokan dilakukan melalui penentuan centroid awal, perhitungan jarak setiap data terhadap centroid, penentuan klaster berdasarkan jarak terdekat, serta pembaruan centroid hingga keanggotaan klaster tidak mengalami perubahan.

Langkah 1. Penentuan Centroid Awal ( Iterasi ke-1)

Centroid awal merupakan titik pusat yang digunakan sebagai acuan dalam menentukan keanggotaan setiap data pada proses klastering. Dalam penelitian ini, centroid awal ditentukan dari data penelitian, kemudian setiap observasi dihitung jaraknya terhadap ketiga centroid tersebut.

**Tabel 3 Nilai Centroid Awal (Iterasi ke-1)**

**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 9601 – 9610 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

| Klaster | Nama Daerah Terpilih | Tahun | X1 (Alpukat) | X2 (Belimbing) |
|---|---|---|---|---|
| K1 (Tinggi) | Tapanuli Selatan | 2022 | 120169 | 1349 |
| K2 (Sedang) | Simalungun | 2022 | 5675 | 2836 |
| K3 (Rendah) | Kota Tanjung Balai | 2022 | 6 | 90 |

Langkah 2. Perhitungan Jarak Euclidean Iterasi ke-1

Perhitungan jarak antara data dan centroid dilakukan menggunakan metode Euclidean Distance. Rumus yang digunakan adalah sebagai berikut:

$$d(xi, uj) = \sqrt{\sum (xi-uj)^2}$$

$$d(\mathrm{ID}_1, C_1) = \sqrt{(122 - 120169)^2 + (48 - 1349)^2}$$
$$d(\mathrm{ID}_1, C_1) = \sqrt{(-120047)^2 + (-1301)^2}$$
$$d(\mathrm{ID}_1, C_1) = 120054{,}05$$
$$d(\mathrm{ID}_1, C_2) = \sqrt{(122 - 5675)^2 + (48 - 2836)^2}$$
$$d(\mathrm{ID}_1, C_2) = \sqrt{(-5553)^2 + (-2788)^2}$$
$$d(\mathrm{ID}_1, C_2) = 6213{,}59$$
$$d(\mathrm{ID}_1, C_3) = \sqrt{(122 - 6)^2 + (48 - 90)^2}$$
$$d(\mathrm{ID}_1, C_3) = \sqrt{(116)^2 + (-42)^2}$$
$$d(\mathrm{ID}_1, C_3) = 123{,}37$$

Jarak terkecil adalah d(ID1) = 123,37 sehingga wilayah Nias dikelompokkan ke dalam Klaster K3.

$$d(\mathrm{ID}_2, C_1) = \sqrt{(11248 - 120169)^2 + (23993 - 1349)^2}$$
$$d(\mathrm{ID}_2, C_1) = \sqrt{(-108921)^2 + (22644)^2}$$
$$d(\mathrm{ID}_2, C_1) = 111249{,}88$$
$$d(\mathrm{ID}_2, C_2) = \sqrt{(11248 - 5675)^2 + (23993 - 2836)^2}$$
$$d(\mathrm{ID}_2, C_2) = \sqrt{(5573)^2 + (21157)^2}$$
$$d(\mathrm{ID}_2, C_2) = 21878{,}69$$
$$d(\mathrm{ID}_2, C_3) = \sqrt{(11248 - 6)^2 + (23993 - 90)^2}$$
$$d(\mathrm{ID}_2, C_3) = \sqrt{(11242)^2 + (23903)^2}$$
$$d(\mathrm{ID}_2, C_3) = 26414{,}69$$

Jarak terkecil adalah d(ID2) = 21878,69 sehingga wilayah Mandailing Natal dikelompokkan ke dalam Klaster K2.

$$d(\mathrm{ID}_3, C_1) = \sqrt{(120169 - 120169)^2 + (1349 - 1349)^2}$$
$$d(\mathrm{ID}_3, C_1) = \sqrt{(0)^2 + (0)^2}$$
$$d(\mathrm{ID}_3, C_1) = 0$$
$$d(\mathrm{ID}_3, C_2) = \sqrt{(120169 - 5675)^2 + (1349 - 2836)^2}$$
$$d(\mathrm{ID}_3, C_2) = \sqrt{(114494)^2 + (-1487)^2}$$
$$d(\mathrm{ID}_3, C_2) = 114503{,}66$$
$$d(\mathrm{ID}_3, C_3) = \sqrt{(120169 - 6)^2 + (1349 - 90)^2}$$
$$d(\mathrm{ID}_3, C_3) = \sqrt{(120163)^2 + (1259)^2}$$
$$d(\mathrm{ID}_3, C_3) = 120169{,}60$$

Jarak terkecil adalah d(ID3) = 0 sehingga wilayah Tapanuli Selatan dikelompokkan ke dalam Klaster K1. Sehingga Hasil keseluruhn Hasil perhitungannya ditampilkan di tabel 4.

**Tabel 4 Hasil Iterasi ke-1**

| Kabupaten/Kota | d(C1) | d(C2) | d(C3) | Min | Klaster |
|---|---|---|---|---|---|
| Nias | 120054.05 | 6213.59 | 123.37 | 123.37 | K3 |
| Mandailing Natal | 111249.88 | 21878.69 | 26414.69 | 21878.69 | K2 |
| Tapanuli Selatan | 0.00 | 114503.66 | 120169.60 | 0.00 | K1 |
| Tapanuli Tengah | 120114.49 | 6285.17 | 99.30 | 99.30 | K3 |
| Tapanuli Utara | 33550.69 | 80975.01 | 86686.59 | 33550.69 | K1 |
| Toba Samosir | 106875.40 | 8134.07 | 13296.25 | 8134.07 | K2 |
| Labuhan Batu | 120169.60 | 6299.05 | 0.00 | 0.00 | K3 |
| Asahan | 114503.66 | 0.00 | 6299.05 | 0.00 | K2 |
| Simalungun | 107105.50 | 7828.56 | 13064.16 | 7828.56 | K2 |
| Dairi | 56000.25 | 58578.69 | 64179.06 | 56000.25 | K1 |
| …… | ……. | …. | .. | | |

| Kabupaten/Kota | d(C1) | d(C2) | d(C3) | Min | Klaster |
|---|---|---|---|---|---|
| …….. | …….. | ……. | … | | |
| Kota Gunungsitoli | 118865.96 | 4879.12 | 1420.98 | 1420.98 | K3 |

Langkah 3. Memperbarui Nilai Centroid

Setelah Iterasi ke-1 selesai, nilai Centroid diperbarui dengan menghitung rata-rata nilai produksi Alpukat dan Belimbing dari seluruh data pada masing-masing klaster. Nilai Centroid baru tersebut digunakan dalam proses pengelompokan pada iterasi berikutnya. Perhitungan Centroid baru menggunakan rumus berikut:

$$uj\ (t+1) = \frac{1}{Nsj}\sum_{j\in j} xj$$

Kluster 1 (Alpukat) = 120169 + 86650 + 64185 + 78924 + 158032 + 91665 + 78286 + 93643 + 79668 + 90344 + 157850 + 63472 /12 = 96907,33

Kluster 1 (Belimbing) = 1349 + 2807 + 0 + 40 + 657 + 2811 + 0 +2851 + 0 + 632 + 0 + 891 / 12 = 1003,17

Kluster 2 (Alpukat) = 11248 +13302 + 5675 + 13069 + 51424 + 10852 + 3513 + 18268 + 13459 + 8117 + 11294 + 25271 + 5985 + 13770 + 41252 + 9881 + 21607 + 9632 + 47728 + 9782 + 11445 + 49118 + 12805 + 5987 + 15775 + 29258 + 6294 + 29602 + 30537 + 8276 + 11446 + 53477 + 8383 + 6612 + 19378 + 22162 + 37024 + 7753 + 9360 / 39 = 18456,95

Kluster 2 (Belimbing) = 23993 + 9 + 2836 + 264 + 0 + 27302 + 8109 + 0 + 0 + 1997 + 23916 + 7 + 2407 + 275 + 0 + 7542 + 0 + 0 + 0 + 160 + 1944 + 23904 + 782 + 13 + 2421 + 256 + 0 + 16653 + 0 + 401 + 1412 + 23871 + 621 + 32 + 2548 + 416 + 0 + 0 + 0 + 1052 / 39 = 4490,85

Kluster 3 (Alpukat) = 122 + 62 + 6 + 710 + 610 + 825 + 324 + 246 + 201 + 0 + 0 + 0 + 0 + 15 + 195 + 700 + 1080 + 1101 + 61 + 23 + 54 + 14 + 2240 + 792 + 981 + 614 + 102 + 45 + 168 + 0 + 50 + 0 + 0 + 9 + 158 + 939 + 936 + 435 + 44 + 40 + 44 + 14 + 2009 + 151 + 3266 + 1935 + 728 + 67 + 0 + 136 + 0 + 51 + 0 + 0 + 30 + 197 + 952 + 984 + 752 + 36 + 162 + 16 + 15 + 1726 + 1157 + 131 + 1441 + 430 + 47 + 220 + 63 + 0 + 30 + 0 + 0 + 65 + 235 + 1401 + 1019 + 903 + 1305 / 81 = 439,75

Kluster 3 (Belimbing) = 48 + 8 + 90 + 1038 + 179 + 647 + 31 + 358 + 31 + 0 + 0 + 0 + 0 + 60 + 25 + 107 + 677 + 330 + 16 + 15 + 0 + 158 + 2539 + 1551 +371 + 474 + 51 + 10 + 49 + 0 + 90 + 50 + 0 + 71 + 17 + 123 + 595 + 488 + 14 + 15 + 0 + 145 + 1545 + 509 + 0 + 504 + 420 + 18 + 0 + 57 + 0 + 111 + 8 + 0 + 69 + 30 + 98 + 513 + 133 + 15 + 58 + 0 + 150 + 1444 + 1019 + 135 + 425 + 139 + 30 + 0 + 44 + 0 + 37 + 96 + 0 + 52 + 53 + 398 + 586 + 85 + 666 / 81 = 245,90

**Langkah 4. Pengelompokan pada Iterasi 2**

Setelah Centroid baru diperoleh, proses klastering dilanjutkan ke Iterasi 2 dengan menghitung kembali jarak setiap data menggunakan metode Euclidean Distance.

**Tabel 5 Hasil Iterasi ke-2**

| Kabupaten/Kota | d(C1) | d(C2) | d(C3) | Min | Klaster |
|---|---|---|---|---|---|
| Nias | 96790.05 | 18865.56 | 374.34 | 374.34 | K3 |
| Mandailing Natal | 88690.78 | 20791.90 | 26091.05 | 20791.90 | K2 |
| Tapanuli Selatan | 23264.24 | 101760.56 | 119734.33 | 23264.24 | K1 |
| Tapanuli Tengah | 96850.45 | 18933.31 | 446.42 | 446.42 | K3 |
| Tapanuli Utara | 10414.73 | 68213.84 | 86248.28 | 10414.73 | K1 |

**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 9601 – 9610 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

| Kabupaten/Kota | d(C1) | d(C2) | d(C3) | Min | Klaster |
|---|---|---|---|---|---|
| Toba Samosir | 83611.24 | 6830.84 | 12864.43 | 6830.84 | K2 |
| Labuhan Batu | 96905.64 | 18968.53 | 460.92 | 460.92 | K3 |
| Asahan | 91250.74 | 12888.63 | 5840.93 | 5840.93 | K3 |
| Simalungun | 83841.59 | 6848.08 | 12629.26 | 6848.08 | K2 |
| Dairi | 32737.71 | 45948.04 | 63745.72 | 32737.71 | K1 |
| ……. | ……. | …. | .. | | |
| ……. | ……. | ……. | … | | |
| Kota Gunungsitoli | 95602.93 | 17573.24 | 961.84 | 961.84 | K3 |

**Langkah 5. Perhitungan hingga Kondisi *Konvergen***

Proses klastering dilakukan secara berulang dengan memperbarui nilai Centroid dan menentukan keanggotaan data berdasarkan jarak terkecil. Iterasi dihentikan ketika tidak terdapat perubahan keanggotaan antar-klaster. Berdasarkan Hasil perhitungan, proses K-Means mencapai konvergensi pada iterasi ke-6, sehingga Hasil pada iterasi tersebut ditetapkan sebagai Hasil akhir pengelompokan.

**Tabel 6. Hasil Iterasi ke-6**

| No | Kabupaten/Kota | Klaster |
|---|---|---|
| 1 | Nias | K3 (Rendah) |
| 2 | Mandailing Natal | K2 (Sedang) |
| 3 | Tapanuli Selatan | K1 (Tinggi) |
| 4 | Tapanuli Tengah | K3 (Rendah) |
| 5 | Tapanuli Utara | K1 (Tinggi) |
| 6 | Toba Samosir | K3 (Rendah) |
| 7 | Labuhan Batu | K3 (Rendah) |
| 8 | Asahan | K3 (Rendah) |
| 9 | Simalungun | K3 (Rendah) |
| 10 | Dairi | K1 (Tinggi) |
| …. | …. | …. |
| …. | …. | …. |
| 132 | Kota Gunungsitoli | K3 (Rendah) |

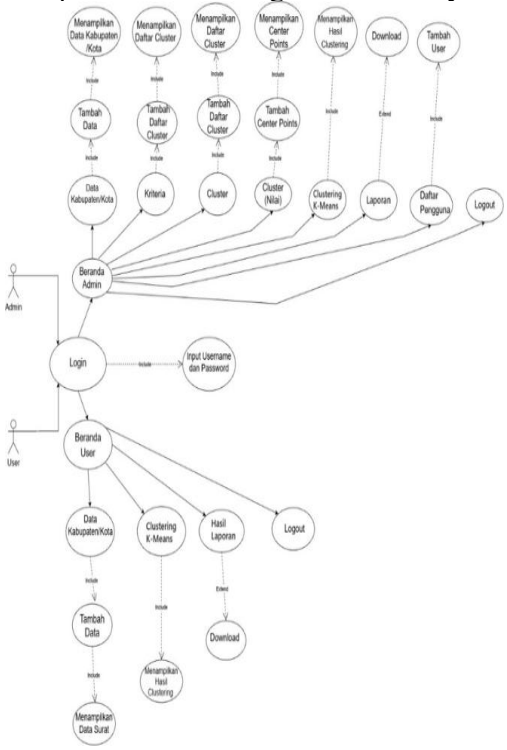
sebanyak 12 data termasuk ke dalam Klaster K1 dengan kategori Tinggi, 20 data termasuk ke dalam Klaster K2 dengan kategori Sedang, dan 100 data termasuk ke dalam Klaster K3 dengan kategori Rendah.

**Pembahasan**

**Analisis Sistem**

Analisis sistem dilakukan untuk memodelkan alur kerja sistem klastering menggunakan pendekatan Unified Modeling Language (UML), yang meliputi Use Case Diagram, Activity Diagram, Sequence Diagram, dan Class Diagram (Mufid & Hanum, Derra Risqi Nurulita, 2023).

Use Case Diagram menggambarkan interaksi antara aktor dan sistem yang dikembangkan. Diagram ini menunjukkan fungsi-fungsi yang dapat diakses oleh setiap aktor sesuai dengan hak aksesnya.




**Gambar 2 Use Case Diagram Sistem**

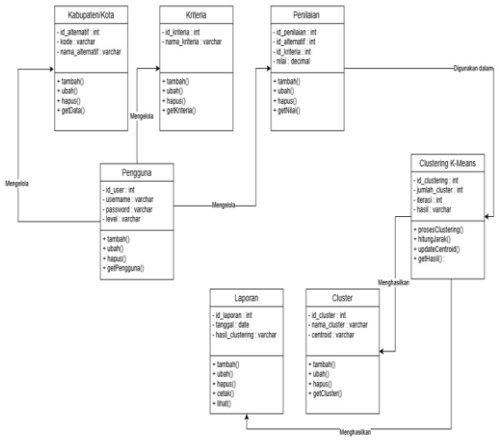
Activity diagram merepresentasikan aliran proses atau aktivitas dalam sebuah sistem yang akan dibangun, mulai dari proses awal, keputusan-keputusan yang

**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 9601 – 9610 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

terjadi di dalam sistem, hingga bagaimana sebuah proses (Septian et al., 2024).

Sequence Diagram pada sistem pengelompokan kabupaten/kota berdasarkan produksi tanaman buah tahunan menggunakan metode K-Means di Provinsi Sumatera Utara menggambarkan urutan interaksi antara Pengguna, Halaman (View), Server PHP, dan Database pada Login serta setiap tahapan sistem (Nafilah et al., 2024).

Class Diagram menggambarkan struktur kelas pada sistem pengelompokan Kabupaten/Kota berdasarkan produksi Alpukat dan Belimbing menggunakan metode K-Means. Diagram ini mencakup kelas pengelolaan data, kriteria, klastering, Hasil pengelompokan, laporan, dan pengguna beserta hubungan antarkelas dalam mendukung proses pengolahan data hingga penyajian Hasil pengelompokan (Putri Milanda Bainamus, Nurfitri Andayani, Chandra Yuliansyah, 2025).

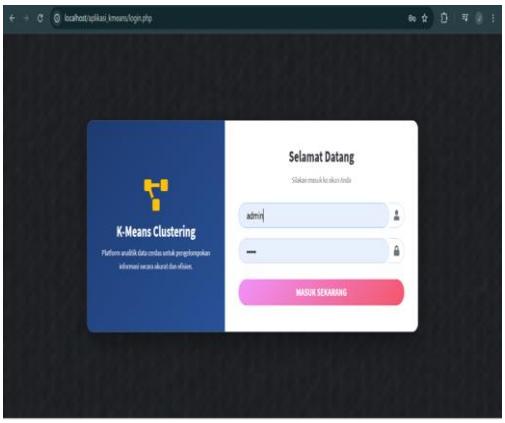



**Gambar 3 Class Diagram Sistem**

**Implementasi Sistem**

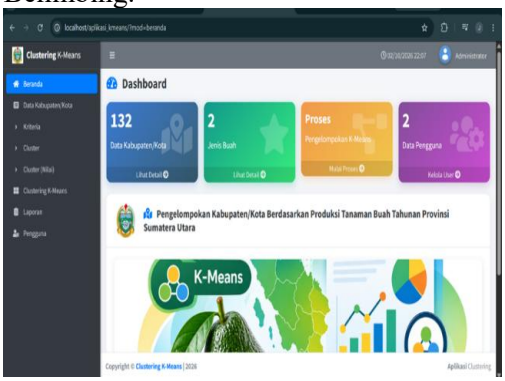
Sistem pengelompokan Kabupaten/Kota berdasarkan produksi tanaman buah tahunan menggunakan metode K-Means di Provinsi Sumatera Utara diimplementasikan sebagai aplikasi berbasis web menggunakan bahasa pemrograman PHP dan basis data MySQL. Sistem ini dirancang untuk mengelola data produksi Alpukat dan Belimbing, melakukan proses klastering, serta menampilkan Hasil pengelompokan dan laporan. Berikut langkah-langkah menjalankan sistem yang dibangun.

1. Halaman login merupakan antarmuka pertama yang ditampilkan saat pengguna mengakses sistem. Pengguna diwajibkan memasukkan username dan password yang telah terdaftar untuk masuk ke dalam sistem dan mengakses fitur yang tersedia.




**Gambar 4 Halaman Login**

Halaman Dashboard merupakan halaman utama yang ditampilkan setelah pengguna berhasil login. Halaman ini menyajikan menu navigasi serta informasi ringkas berupa 132 data Kabupaten/Kota, 2 jenis buah, proses K-Means, dan 2 pengguna. Dashboard juga menampilkan ilustrasi, peta Provinsi Sumatera Utara, dan grafik yang mendukung visualisasi pengelompokan data produksi Alpukat dan Belimbing.

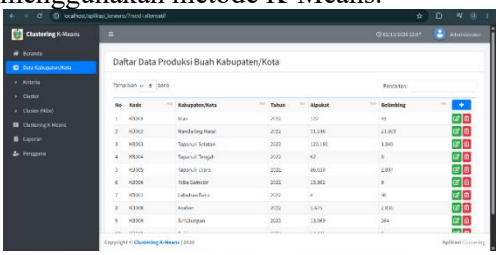



**Gambar 5 Halaman Dashboard**

Halaman Data Kabupaten/Kota digunakan untuk menampilkan dan mengelola data produksi Alpukat dan Belimbing berdasarkan Kabupaten/Kota dan tahun. Pengguna dapat menambah, mengubah, dan menghapus data yang tersimpan dalam database. Data tersebut

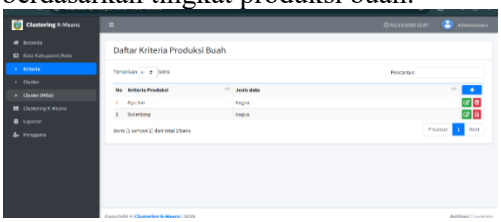
**Journal of Science and Social Research** ISSN 2615 – 4307 (Print)
Oct 2026, IX (5): 9601 – 9610 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

selanjutnya digunakan sebagai masukan dalam proses pengelompokan menggunakan metode K-Means.



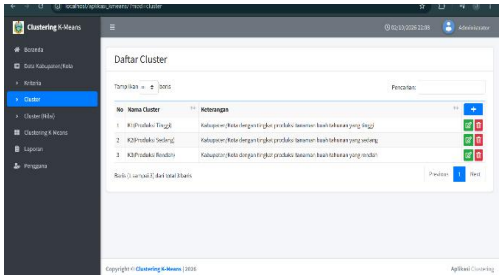

**Gambar 6 Halaman Data Kabupaten/Kota**

Halaman Kriteria menampilkan dua variabel yang digunakan dalam proses klastering, yaitu produksi Alpukat dan Belimbing. Kedua kriteria tersebut menjadi dasar perhitungan jarak antar data dalam metode K-Means untuk mengelompokkan Kabupaten/Kota berdasarkan tingkat produksi buah.



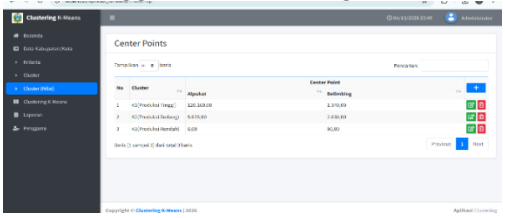

**Gambar 7 Halaman Kriteria**

Halaman Klaster menampilkan tiga kelompok pengelompokan, yaitu K1 (Tinggi), K2 (Sedang), dan K3 (Rendah). Ketiga kelompok tersebut digunakan untuk mengategorikan Kabupaten/Kota berdasarkan tingkat produksi Alpukat dan Belimbing menggunakan metode K-Means.



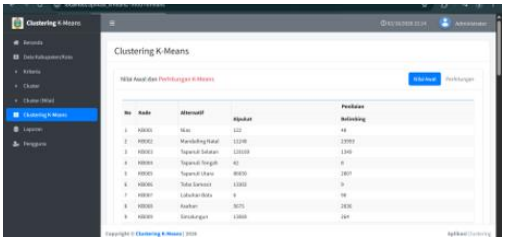

**Gambar 8 Halaman Klaster**

Halaman Klaster Nilai menampilkan nilai Centroid untuk setiap kelompok K1, K2, dan K3. Nilai tersebut digunakan sebagai titik pusat dalam perhitungan jarak data produksi Alpukat dan Belimbing serta diperbarui pada setiap iterasi hingga Hasil klastering stabil.



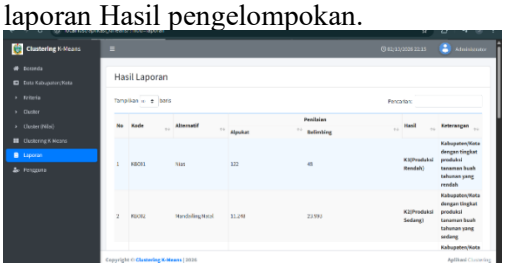

**Gambar 9 Halaman Klaster Nilai**

Halaman Klastering K-Means digunakan untuk menjalankan proses pengelompokan data Kabupaten/Kota berdasarkan produksi Alpukat dan Belimbing. Sistem menghitung jarak menggunakan Euclidean Distance, menentukan klaster berdasarkan jarak terdekat, dan memperbarui centroid hingga diperoleh Hasil pengelompokan yang stabil.




**Gambar 10 Halaman Klastering K-Means**

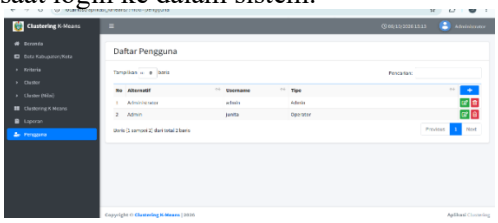
Halaman Laporan menampilkan Hasil klastering Kabupaten/Kota yang mencakup kode dan nama wilayah, nilai produksi Alpukat dan Belimbing, serta kategori klaster berdasarkan tingkat produksi. Data disajikan dalam tabel dengan navigasi halaman dan dilengkapi fitur Export PDF untuk mengunduh laporan Hasil pengelompokan.




**Gambar 11 Halaman Laporan**

Halaman Pengguna digunakan untuk mengelola akun yang memiliki akses ke sistem. Pengguna dengan hak akses yang sesuai dapat menambah,

mengubah, dan menghapus akun. Data akun digunakan untuk proses autentikasi saat login ke dalam sistem.




**Gambar 12 Halaman Pengguna**

## SIMPULAN

Penelitian ini menerapkan metode K-Means untuk mengelompokkan kabupaten/kota di Provinsi Sumatera Utara berdasarkan produksi alpukat dan belimbing pada periode 2022–2025. Data penelitian terdiri atas 132 pengamatan dari 33 kabupaten/kota dengan dua variabel produksi. Proses perhitungan mencapai kondisi konvergen pada iterasi ke-6 dan menghasilkan tiga klaster, yaitu K1 (tinggi) sebanyak 12 pengamatan, K2 (sedang) sebanyak 20 pengamatan, dan K3 (rendah) sebanyak 100 pengamatan. Klaster rendah memiliki jumlah anggota terbesar.

Metode tersebut diimplementasikan dalam aplikasi berbasis web menggunakan PHP dan MySQL untuk mendukung pengelolaan data, proses klastering, penyajian hasil, dan pembuatan laporan. Hasil pengelompokan dapat menjadi informasi pendukung untuk memahami persebaran karakteristik produksi alpukat dan belimbing. Penelitian selanjutnya disarankan menambahkan variabel yang relevan serta melakukan evaluasi kualitas klaster atau membandingkan K-Means dengan metode klastering lain.

## DAFTAR PUSTAKA

Annisa Wulandari, I. (2025). Penerapan Metode K-Means Clustering untuk pemetaan Perbaikan Jalan di Kota Bogor. *Jurnal Informatika Teknologi Dan Sains (JINTEKS)*, *7*(1), 107–116.

Harun Mukhtar, T. M. S. (2024). Analisis Kesuburan Pertanian Melalui Irigasi Dengan Menggunakan Metode K-Means Clustering. *Jurnal Software Engineering and Information System (SEIS)*, *4*(2), 102–107.

Mufid, N. A., & Hanum, Derra Risqi Nurulita, A. H. S. (2023). Clustering Kabupaten / Kota di Jawa Tengah Tahun 2022 berdasarkan Jumlah Kasus Kemunculan Penyakit dengan Algoritma K-Means. *Jurnal Pendidikan Matematika*, *1*(1), 1–14.

Nafilah, M., Rahaningsih, N., & Danar Dana, R. (2024). Perbandingan Algoritma K-Means Dan K-Medoids Untuk Pengelompokan Hasil Pertanian Di Kabupaten Cirebon. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 223–229.

Nilam Kumala, R. D. D. (2023). Penerapan Algoritma K-Means Clustering Berdasarkan Hasil Penjualan Produk Inez Cosmetic di Toko Rosalia. *Jurnal Teknik Elektro Dan Informatika*, *18*(1), 50–56.

Putri Milanda Bainamus, Nurfitri Andayani, Chandra Yuliansyah, R. A. (2025). Rancangan Bangun Sistem Informasi Akademik Berbasis Web Pada SDN 07 Bermani Ilir Kabupaten Kepahiang. *Jurnal Ilmiah Teknik Dan Ilmu Komputer*, *4*(2), 60–68.

Risbuwono Heru Cokro, D. A. A. (2025). Segmentasi Provinsi Penghasil Padi Di Indonesia Berdasarkan Karakteristik Produksi Tahun 2024 Menggunakan Clustering K-Means. *Jurnal Multidisiplin Saintek*, *8*(10), 1–9.

Septian, R., Achmad Fauzi, & Milli Alfhi Syari. (2024). Design and Build Automatic Fan Control Based Internet of Things. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, *4*(1), 339–347.