

PENGELOMPOKKAN TINGKAT HARGA CABAI RAWIT BERDASARKAN PROVINSI DI INDONESIA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING

(Grouping Chillies Price Levels Based On Province In Indonesia Using K-Means Clustering Algorithm)

Citra Fathia Palembang¹, Septianti Permatasari Palembang²

^{1,2} Universitas Pattimura

Jl. Ir. M. Putuhena, Kampus Poka, Kota Ambon-Maluku, Indonesia.

e-mail: citra.palembang@fmipa.unpatti.ac.id^{1*}, septiantipermatasari@gmail.com²

Abstrak: Sebagai salah satu pangan pokok masyarakat yang digunakan untuk bumbu masak, bahkan obat-obatan, harga komoditas cabai selalu berfluktuasi dan mengalami kenaikan harga yang signifikan. Hal ini dikarenakan karakteristik cabai yang bertumpu pada musim dan cuaca. Harga rata-rata secara nasional tercatat Rp.41.244,-/Kg. Tingginya permintaan kebutuhan akan cabai rawit dan kurangnya stok menjadi pemicu mahalannya harga cabai rawit di pasaran. Pada penelitian ini membahas tentang Pengelompokan tingkat harga cabai rawit di tahun 2020-2021 berdasarkan provinsi di Indonesia menggunakan algoritma K-Means clustering. Data komoditas cabai yang diperoleh dari situs resmi Pusat Informasi Harga Pangan Nasional Strategis (PIHPS) dibagi menjadi tiga cluster, yaitu wilayah dengan cluster harga cabai tinggi, cluster dengan harga sedang, dan rendah. Selain melakukan perhitungan secara manual, peneliti juga menggunakan tool RapidMiner dalam pemrosesan algoritma K-Means Clustering untuk dapat membandingkan hasil perhitungan. Hasil pengelompokan diperoleh Cluster harga cabai tertinggi berada di 10 provinsi yaitu: Kepulauan Babel, DKI Jakarta, NTT, Kalbar, Kalteng, Kaltara, Maluku, Maluku Utara, Papua dan Papua Barat. Hasil dari pengelompokan harga cabai tersebut dapat dijadikan sebagai bahan evaluasi atau dasar pertimbangan pihak-pihak terkait membuat kebijakan dalam hal menjamin ketersediaan, mengurangi disparitas harga, dan menjaga kestabilan harga baik pada saat harga tinggi maupun pada saat harga jatuh.

Kata Kunci: K-Means, Clustering, Cabai Rawit.

Abstract: As one of the people's staple foods used for cooking spices, even medicines, the price of chillies always fluctuate and have significant price increases. It's because the characteristics of chillies are based on the season and weather. The average price was IDR 41.244/Kg. The high demand for and lack of stock are the triggers for the high price of chillies in the market. This study discusses the grouping of chillies price levels in 2020-2021 based on provinces in Indonesia using the K-Means clustering algorithm. Chili commodity data obtained from the official website of Pusat Informasi Harga Pangan Nasional Strategis (PIHPS) is divided into three clusters, there are the areas with high chili price clusters, medium price clusters, and low price clusters. In addition to doing calculations manually, researchers also use the RapidMiner tool in processing the K-Means Clustering algorithm to be able to compare the calculation results. The results of the grouping obtained that the highest chili price clusters were in 10 provinces, namely: Kepulauan Babel, DKI Jakarta, NTT, Kalbar, Kalteng, Kaltara, Maluku, Maluku Utara, Papua dan Papua Barat. The results of the grouping of chili prices can be used as evaluation material or the basis for consideration of related parties to make policies in terms of ensuring availability, reducing price disparities, and maintaining price stability both when prices are high and when prices fall.

Keywords: K-Means, Clustering, Chili.

1. PENDAHULUAN

Cabai merupakan produk hortikultura yang banyak dikonsumsi masyarakat. Cabai dipergunakan untuk keperluan rumah tangga sebagai rempah dan bumbu masakan, juga digunakan dalam berbagai keperluan industri diantaranya, industri bumbu masakan, industri makanan, industri obat-obatan, dan

industri kosmetik. Cabai merupakan salah satu komoditas strategis yang ditetapkan sebagai bahan pangan pokok yang tidak tergantikan dan menjadi perhatian serius pemerintah karena memberikan andil yang cukup signifikan dalam menentukan inflasi. Hal ini tidak terlepas dari karakteristik dari komoditas cabai itu sendiri yang sangat bertumpu pada cuaca dan musim.

Berdasarkan data dari Pusat Informasi Harga Pangan Strategis (PIHPS) Nasional, per Juli 2021, secara rata-rata harga cabai mengalami kenaikan yang cukup signifikan. Kenaikan harga cabai rawit per kg setara dengan harga daging ayam per ekor mencapai Rp.67.500,-/Kg pada akhir Juli 2021 di Provinsi Maluku. Tidak hanya di Maluku, kenaikan harga cabai rawit hampir di seluruh provinsi juga ikut mengalami hal tersebut. Harga rata-rata secara nasional yakni sebesar Rp. 41.244,-/Kg.

Musim panen dan musim tanam serta pengaruh iklim dan cuaca sangat mempengaruhi kenaikan harga cabai di pasaran. Bila dibandingkan dengan harga di daerah konsumen, harga cabai di daerah produsen lebih rendah karena adanya faktor angkutan, rendahnya daya tahan cabai, dan daya beli masyarakat yang rendah. Oleh karena adanya musim produksi cabai di masing-masing daerah, harga pasar lokal menjadi sangat berfluktuasi, yang disebabkan jatuhnya harga pada saat panen raya [1].

Selain itu, adanya pandemi Covid-19 yang muncul sejak Maret 2020 ikut menyebabkan penyerapan produksi cabai turun 90 persen¹. Hal ini sebab dilakukannya pembatasan maupun penutupan hotel, restoran, kafe, dan pasar tradisional di masa pandemi untuk menekan transmisi virus corona, akhirnya pasokan jadi melebihi permintaan. Harga jual yang rendah membuat petani kekurangan modal untuk menanam kembali. Di sisi lain, banyak petani yang jadi enggan menanam cabai kembali karena harga jualnya tak sesuai skala ekonomi. Selain itu, kondisi cuaca ekstrem dengan curah hujan yang tinggi di sejumlah wilayah, turut berpengaruh pada produksi cabai.

Kurangnya pasokan cabai merupakan imbas dari kerugian besar-besaran yang dialami petani cabai, dimana saat pasokan melimpah namun harga cabai anjlok. Rendahnya produksi cabai dalam negeri membuat pasokan di pasar tak bisa memenuhi tingginya permintaan masyarakat menyebabkan permintaan yang lebih tinggi membuat harga cabai terus melambung.

Clustering merupakan metode yang cara kerjanya mencari dan mengelompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya dengan Ciri khas mempunyai sifat tanpa arahan (*unsupervised*), artinya teknik ini dapat diterapkan tanpa perlunya data *training* dan instruktur serta tidak diperlukannya target output[2]. Salah satu Metode clustering yang efisien dan cepat adalah Algoritma K-Means yang bekerja pada atribut angka atau numerik dengan mengandalkan teknik Clustering berbasis jarak dengan membagi data ke dalam beberapa cluster[3] menggunakan data yang mirip yang dikelompokkan pada cluster yang sama [4]. Penelitian sebelumnya yang menggunakan Metode Clustering Algoritma K-Means dalam menyelesaikan beberapa persoalan diantaranya: Penelitian yang mengelompokkan daerah berdasarkan potensi produksi karet diukur berdasarkan perhitungan nilai *purity measure* dari masing-masing cluster [5], penelitian berikutnya mengelompokkan data penjualan alat-alat bangunan sehingga dapat membantu pihak manajemen dalam menentukan prioritas pembelian stok alat-alat bangunan yang laris terjual dan mengurangi pembelian stok alat-alat yang kurang laris [6], Penelitian lainnya mengelompokkan rata-rata harga eceran beras agar kota yang masuk ke dalam cluster tingkat harga tinggi dapat menormalisasikan harga eceran beras di tiap-tiap kota tersebut [7], juga penelitian dalam mengelompokkan data imunisasi campak pada balita berdasarkan provinsi yang hasilnya provinsi yang masuk dalam provinsi cluster tertinggi agar mendapat perhatian lebih dari pemerintah dan dapat meningkatkan sosialisasi imunisasi di provinsi tersebut[8]. Peneliti menggunakan algoritma K-means karena merupakan metode clustering paling sederhana dan umum, memiliki kemampuan pengelompokkan data dalam jumlah besar dengan waktu komputasi yang relatif cepat dan efisien [7]. Beberapa penelitian sebelumnya hanya menggunakan salah satu teknik yaitu mengelompokkan data secara manual menggunakan rumus atau hanya menggunakan *tool*, maka dalam penelitian ini, peneliti menggunakan keduanya yaitu perhitungan secara manual dengan Algoritma K-Means juga menggunakan *tool* RapidMiner untuk mengelompokkan tingkat harga cabai yang tinggi, sedang dan rendah berdasarkan provinsi, Sehingga tingkat akurasi pengelompokkan harga berdasarkan provinsi lebih pasti.

pengelompokan harga cabai tersebut nantinya dapat dijadikan sebagai bahan evaluasi atau dasar pertimbangan oleh pihak-pihak yang terkait dalam membuat kebijakan dalam hal menjamin ketersediaan, mengurangi disparitas harga, dan menjaga kestabilan harga baik pada saat harga tinggi maupun pada saat harga jatuh.

2. METODOLOGI

2.1 Tipe Penelitian

Tipe Penelitian ini merupakan penelitian kuantitatif yaitu teknik mengelompokkan kenaikan harga cabai menggunakan algoritma clustering K-Means dengan *tool* RapidMiner. Data yang digunakan dalam penelitian ini yaitu data laporan bulanan harga cabai rawit seluruh provinsi di Indonesia selama 1 tahun mulai dari Juli 2020 hingga Juli 2021, yang merupakan data sekunder dari situs resmi Pusat Informasi Harga Pangan Strategis (PHIPS) Nasional, <https://hargapangan.id/>.

2.2 Tempat dan Waktu Penelitian

Lokasi Penelitian dilakukan di Program Studi Matematika Fakultas MIPA Universitas Pattimura Ambon.

2.3 Variabel Penelitian

Variabel penelitian ini terdiri dari variabel kategori, yaitu: Kategori cluster tingkat harga cabai tinggi, tingkat harga cabai sedang dan tingkat harga cabai rendah, yang kemudian akan dipakai cara klasifikasi data mining algoritma K-Means dalam menentukan kategori cluster tingkatan harga cabai.

3. HASIL DAN PEMBAHASAN

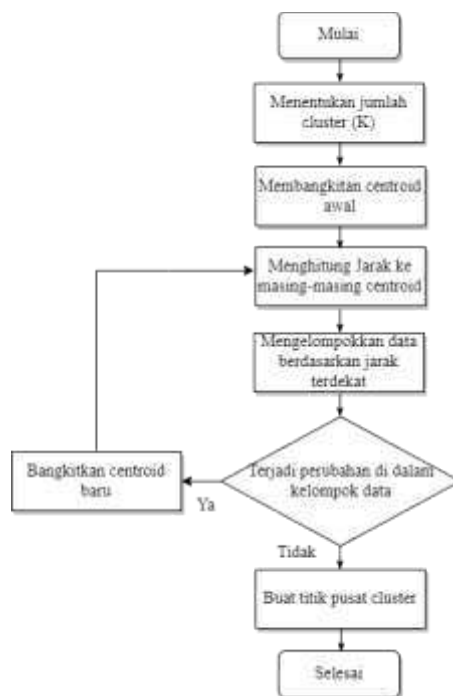
Sebelum memulai proses pengolahan data menggunakan algoritma K-Means, dilakukan *pre processing data*, dimana data laporan bulanan selama 1 tahun mulai dari bulan Juli 2020 sampai dengan Juli 2021 dicari harga rata-rata per provinsi terlebih dahulu, yang ditunjukkan pada tabel 1 untuk dapat digunakan dalam proses clustering.

Tabel 1. Harga Rata-rata Cabai Per Provinsi tahun 2020-2021

| No | Provinsi | Rata-rata (Rp) | No | Provinsi | Rata-rata (Rp) |
|----|---------------------------|----------------|----|---------------------|----------------|
| 1 | Aceh | 36.096 | 18 | Nusa Tenggara Barat | 32.581 |
| 2 | Sumatera Utara | 33.715 | 19 | Nusa Tenggara Timur | 55.438 |
| 3 | Sumatera Barat | 38.142 | 20 | Kalimantan Barat | 47.673 |
| 4 | Riau | 38.804 | 21 | Kalimantan Selatan | 42054 |
| 5 | Kepulauan Riau | 43.227 | 22 | Kalimantan Tengah | 61308 |
| 6 | Jambi | 31.831 | 23 | Kalimantan Timur | 43519 |
| 7 | Bengkulu | 33.350 | 24 | Kalimantan Utara | 53581 |
| 8 | Sumatera Selatan | 39.650 | 25 | Gorontalo | 40538 |
| 9 | Kepulauan Bangka Belitung | 58.273 | 26 | Sulawesi Selatan | 28673 |
| 10 | Lampung | 34.846 | 27 | Sulawesi Tenggara | 38304 |
| 11 | Banten | 40.750 | 28 | Sulawesi Tengah | 31623 |
| 12 | Jawa Barat | 39.942 | 29 | Sulawesi Utara | 37019 |
| 13 | Dki Jakarta | 47.338 | 30 | Sulawesi Barat | 29473 |
| 14 | Jawa Tengah | 31.685 | 31 | Maluku | 62577 |
| 15 | Di Yogyakarta | 33.381 | 32 | Maluku Utara | 47365 |
| 16 | Jawa Timur | 30.908 | 33 | Papua | 59727 |
| 17 | Bali | 28.281 | 34 | Papua Barat | 50615 |

3.1 Implementasi Algoritma K-Means secara manual menggunakan Ms.Excel

Secara umum tahapan clustering dengan menggunakan Algoritma K-Means menggunakan cara manual dapat dilakukan sebagai berikut:



Gambar 1. Alur Algoritma K-Means

1. Menentukan jumlah cluster (K)
 Misal pada penelitian ini jumlah cluster yang akan dibuat sebanyak 3 cluster yaitu cluster dengan harga tertinggi, cluster dengan harga sedang dan cluster dengan harga terendah.
2. Menentukan centroid
 Membangkitkan centroid awal yang dipilih acak secara manual dari data yang digunakan [9]. Jumlah centroid sebanyak yang ditentukan di awal menjadi awal pusat cluster, Penentuan centroid akan berpengaruh ketika menentukan jumlah cluster yang digunakan [10].

Tabel 2. Data Centroid awal dipilih acak

| Centroid 0 | Centroid 1 | Centroid 2 |
|------------|------------|------------|
| 55438 | 36096 | 33715 |

Pada tahap iterasi, digunakan rumus berikut dalam menentukan nilai centroid berikutnya:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \quad \dots \quad (1)$$

Keterangan:

- V_{ij} : Centroid rata-rata cluster ke- i untuk variabel ke- j
- N_i : Jumlah anggota cluster ke- i
- k Indeks dari cluster j : Indeks dari variabel
- X_{kj} : nilai data ke- k variabel ke- j untuk cluster tersebut [3].

3. Menghitung jarak objek dari centroid, menggunakan rumus Euclidean Distance:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots (2)$$

Keterangan:

d : Euclidean Distance

x, y : Koordinat Objek

i : Banyaknya Objek

Tabel 3. Hasil perhitungan jarak objek ke centroid

| Perhitungan Jarak Objek Ke Centroid | | | | |
|-------------------------------------|------------------------------|--------------------|--------------------|--------------------|
| No | Provinsi | C0 (Centroid 0) | C1 (Centroid 1) | C2 (Centroid 2) |
| 1 | Aceh | 19342 | 0 | 2381 |
| 2 | Sumatera Utara | 21723 | 2381 | 0 |
| 3 | Sumatera Barat | 17296 | 2046 | 4427 |
| 4 | Riau | 16634 | 2708 | 5089 |
| 5 | Kepulauan Riau | 12211 | 7131 | 9512 |
| 6 | Jambi | 23607 | 4265 | 1884 |
| 7 | Bengkulu | 22088 | 2746 | 365 |
| 8 | Sumatera Selatan | 15788 | 3554 | 5935 |
| 9 | Kepulauan Bangka Belitung | 2835 | 22177 | 24558 |
| 10 | Lampung | 20592 | 1250 | 1131 |
| 11 | Banten | 14688 | 4654 | 7035 |
| 12 | Jawa Barat | 15496 | 3846 | 6227 |
| 13 | Dki Jakarta | 8100 | 11242 | 13623 |
| 14 | Jawa Tengah | 23753 | 4411 | 2030 |
| 15 | Di Yogyakarta | 22057 | 2715 | 334 |
| 16 | Jawa Timur | 24530 | 5188 | 2807 |
| 17 | Bali | 27157 | 7815 | 5434 |
| 18 | Nusa Tenggara Barat | 22857 | 3515 | 1134 |
| 19 | Nusa Tenggara Timur | 0 | 19342 | 21723 |
| 20 | Kalimantan Barat | 7765 | 11577 | 13958 |
| 21 | Kalimantan Selatan | 13384 | 5958 | 8339 |
| 22 | Kalimantan Tengah | 5870 | 25212 | 27593 |
| 23 | Kalimantan Timur | 11919 | 7423 | 9804 |
| 24 | Kalimantan Utara | 1857 | 17485 | 19866 |
| 25 | Gorontalo | 14900 | 4442 | 6823 |
| 26 | Sulawesi Selatan | 26765 | 7423 | 5042 |
| 27 | Sulawesi Tenggara | 17134 | 2208 | 4589 |
| 28 | Sulawesi Tengah | 23815 | 4473 | 2092 |
| 29 | Sulawesi Utara | 18419 | 923 | 3304 |
| 30 | Sulawesi Barat | 25965 | 6623 | 4242 |
| 31 | Maluku | 7139 | 26481 | 28862 |
| 32 | Maluku Utara | 8073 | 11269 | 13650 |
| 33 | Papua | 4289 | 23631 | 26012 |
| 34 | Papua Barat | 4823 | 14519 | 16900 |

4. Mengelompokkan Data berdasarkan jarak terdekat/terpendek

Dari tabel 3 dan tabel 4 hasil perhitungan jarak objek ke centroid, maka selanjutnya mencari nilai minimal dari ketiga centroid yang dihasilkan. Maka tabel pengelompokkan data berdasarkan jarak, iterasi pertama adalah sebagai berikut:

Tabel 4. Pengelompokkan data berdasarkan jarak (Iterasi 1)

| Perhitungan Jarak Objek Ke Centroid | | | |
|-------------------------------------|---------------------------|---|-------------------------|
| No | Provinsi | Jarak Terpendek (Nilai Centroid terendah dari ketiga cluster) | Kelompok Data Iterasi 1 |
| 1 | Aceh | 0 | Centroid 1 |
| 2 | Sumatera Utara | 0 | Centroid 2 |
| 3 | Sumatera Barat | 2046 | Centroid 1 |
| 4 | Riau | 2708 | Centroid 1 |
| 5 | Kepulauan Riau | 7131 | Centroid 1 |
| 6 | Jambi | 1884 | Centroid 2 |
| 7 | Bengkulu | 365 | Centroid 2 |
| 8 | Sumatera Selatan | 3554 | Centroid 1 |
| 9 | Kepulauan Bangka Belitung | 2835 | Centroid 0 |
| 10 | Lampung | 1131 | Centroid 2 |
| 11 | Banten | 4654 | Centroid 1 |
| 12 | Jawa Barat | 3846 | Centroid 1 |
| 13 | Dki Jakarta | 8100 | Centroid 0 |
| 14 | Jawa Tengah | 2030 | Centroid 2 |
| 15 | Di Yogyakarta | 334 | Centroid 2 |
| 16 | Jawa Timur | 2807 | Centroid 2 |
| 17 | Bali | 5434 | Centroid 2 |
| 18 | Nusa Tenggara Barat | 1134 | Centroid 2 |
| 19 | Nusa Tenggara Timur | 0 | Centroid 0 |
| 20 | Kalimantan Barat | 7765 | Centroid 0 |
| 21 | Kalimantan Selatan | 5958 | Centroid 1 |
| 22 | Kalimantan Tengah | 5870 | Centroid 0 |
| 23 | Kalimantan Timur | 7423 | Centroid 1 |
| 24 | Kalimantan Utara | 1857 | Centroid 0 |
| 25 | Gorontalo | 4442 | Centroid 1 |
| 26 | Sulawesi Selatan | 5042 | Centroid 2 |
| 27 | Sulawesi Tenggara | 2208 | Centroid 1 |
| 28 | Sulawesi Tengah | 2092 | Centroid 2 |
| 29 | Sulawesi Utara | 923 | Centroid 1 |
| 30 | Sulawesi Barat | 4242 | Centroid 2 |
| 31 | Maluku | 7139 | Centroid 0 |
| 32 | Maluku Utara | 8073 | Centroid 0 |
| 33 | Papua | 4289 | Centroid 0 |
| 34 | Papua Barat | 4823 | Centroid 0 |

Dari hasil perhitungan jarak pada iterasi pertama menghasilkan pembentukan cluster baru

Tabel 5. Pembentukan Cluster Baru

| Centroid 0 | Centroid 1 | Centroid 2 |
|------------|------------|------------|
| 54389,50 | 36772,69 | 31695,58 |

5. Jika terjadi perubahan di dalam kelompok data, maka lakukan perhitungan ulang. Karena terjadi perubahan di dalam kelompok data setelah iterasi ke-1, maka dilakukan perhitungan ulang data menggunakan nilai centroid baru dan pengelompokan data berdasarkan jarak terpendek, iterasi ke-2 yang digabung menjadi satu tabel seperti berikut :

Tabel 6. Iterasi ke-2

| Perhitungan Jarak Objek Ke Centroid | | | | | | Kelompok Data 2 |
|-------------------------------------|---------------------------|--------------------|--------------------|--------------------|-----------------|-----------------|
| No | Provinsi | C0 (Centroid 1) | C1 (Centroid 2) | C2 (Centroid 3) | Jarak Terpendek | |
| 1 | Aceh | 18293,50 | 676,69 | 4400,42 | 676,69 | Centroid 1 |
| 2 | Sumatera Utara | 20674,50 | 3057,69 | 2019,42 | 2019,42 | Centroid 2 |
| 3 | Sumatera Barat | 16247,50 | 1369,31 | 6446,42 | 1369,31 | Centroid 1 |
| 4 | Riau | 15585,50 | 2031,31 | 7108,42 | 2031,31 | Centroid 1 |
| 5 | Kepulauan Riau | 11162,50 | 6454,31 | 11531,42 | 6454,31 | Centroid 1 |
| 6 | Jambi | 22558,50 | 4941,69 | 135,42 | 135,42 | Centroid 2 |
| 7 | Bengkulu | 21039,50 | 3422,69 | 1654,42 | 1654,42 | Centroid 2 |
| 8 | Sumatera Selatan | 14739,50 | 2877,31 | 7954,42 | 2877,31 | Centroid 1 |
| 9 | Kepulauan Bangka Belitung | 3883,50 | 21500,31 | 26577,42 | 3883,50 | Centroid 0 |
| 10 | Lampung | 19543,50 | 1926,69 | 3150,42 | 1926,69 | Centroid 1 |
| 11 | Banten | 13639,50 | 3977,31 | 9054,42 | 3977,31 | Centroid 1 |
| 12 | Jawa Barat | 14447,50 | 3169,31 | 8246,42 | 3169,31 | Centroid 1 |
| 13 | Dki Jakarta | 7051,50 | 10565,31 | 15642,42 | 7051,50 | Centroid 0 |
| 14 | Jawa Tengah | 22704,50 | 5087,69 | 10,58 | 10,58 | Centroid 2 |
| 15 | Di Yogyakarta | 21008,50 | 3391,69 | 1685,42 | 1685,42 | Centroid 2 |
| 16 | Jawa Timur | 23481,50 | 5864,69 | 787,58 | 787,58 | Centroid 2 |
| 17 | Bali | 26108,50 | 8491,69 | 3414,58 | 3414,58 | Centroid 2 |
| 18 | Nusa Tenggara Barat | 21808,50 | 4191,69 | 885,42 | 885,42 | Centroid 2 |
| 19 | Nusa Tenggara Timur | 1048,50 | 18665,31 | 23742,42 | 1048,50 | Centroid 0 |
| 20 | Kalimantan Barat | 6716,50 | 10900,31 | 15977,42 | 6716,50 | Centroid 0 |
| 21 | Kalimantan Selatan | 12335,50 | 5281,31 | 10358,42 | 5281,31 | Centroid 1 |
| 22 | Kalimantan Tengah | 6918,50 | 24535,31 | 29612,42 | 6918,50 | Centroid 0 |
| 23 | Kalimantan Timur | 10870,50 | 6746,31 | 11823,42 | 6746,31 | Centroid 1 |
| 24 | Kalimantan Utara | 808,50 | 16808,31 | 21885,42 | 808,50 | Centroid 0 |
| 25 | Gorontalo | 13851,50 | 3765,31 | 8842,42 | 3765,31 | Centroid 1 |
| 26 | Sulawesi Selatan | 25716,50 | 8099,69 | 3022,58 | 3022,58 | Centroid 2 |
| 27 | Sulawesi Tenggara | 16085,50 | 1531,31 | 6608,42 | 1531,31 | Centroid 1 |
| 28 | Sulawesi Tengah | 22766,50 | 5149,69 | 72,58 | 72,58 | Centroid 2 |
| 29 | Sulawesi Utara | 17370,50 | 246,31 | 5323,42 | 246,31 | Centroid 1 |
| 30 | Sulawesi Barat | 24916,50 | 7299,69 | 2222,58 | 2222,58 | C2 |

| | | | | | | |
|----|--------------|---------|----------|----------|---------|----|
| 31 | Maluku | 8187,50 | 25804,31 | 30881,42 | 8187,50 | C0 |
| 32 | Maluku Utara | 7024,50 | 10592,31 | 15669,42 | 7024,50 | C0 |
| 33 | Papua | 5337,50 | 22954,31 | 28031,42 | 5337,50 | C0 |
| 34 | Papua Barat | 3774,50 | 13842,31 | 18919,42 | 3774,50 | C0 |

Setelah dilakukan iterasi ke-2, hasil pada kelompok data masih berubah, maka lakukan lagi iterasi ke-3 sampai kepada iterasi yang menemukan kelompok data tidak lagi berubah atau sama, maka dibuat titik pusat cluster.

Tabel 7. Iterasi ke-3

| Perhitungan Jarak Objek Ke Centroid | | | | | | Kelompok Data 3 |
|-------------------------------------|---------------------------|--------------------|--------------------|--------------------|-----------------|-----------------|
| No | Provinsi | C0 (Centroid 1) | C1 (Centroid 2) | C2 (Centroid 3) | Jarak Terpendek | |
| 1 | Aceh | 18293,50 | 3357,15 | 4686,82 | 3357,15 | Centroid 1 |
| 2 | Sumatera Utara | 20674,50 | 5738,15 | 2305,82 | 2305,82 | Centroid 2 |
| 3 | Sumatera Barat | 16247,50 | 1311,15 | 6732,82 | 1311,15 | Centroid 1 |
| 4 | Riau | 15585,50 | 649,15 | 7394,82 | 649,15 | Centroid 1 |
| 5 | Kepulauan Riau | 11162,50 | 3773,85 | 11817,82 | 3773,85 | Centroid 1 |
| 6 | Jambi | 22558,50 | 7622,15 | 421,82 | 421,82 | Centroid 2 |
| 7 | Bengkulu | 21039,50 | 6103,15 | 1940,82 | 1940,82 | Centroid 2 |
| 8 | Sumatera Selatan | 14739,50 | 196,85 | 8240,82 | 196,85 | Centroid 1 |
| 9 | Kepulauan Bangka Belitung | 3883,50 | 18819,85 | 26863,82 | 3883,50 | Centroid 0 |
| 10 | Lampung | 19543,50 | 4607,15 | 3436,82 | 3436,82 | Centroid 2 |
| 11 | Banten | 13639,50 | 1296,85 | 9340,82 | 1296,85 | Centroid 1 |
| 12 | Jawa Barat | 14447,50 | 488,85 | 8532,82 | 488,85 | Centroid 1 |
| 13 | Dki Jakarta | 7051,50 | 7884,85 | 15928,82 | 7051,50 | Centroid 0 |
| 14 | Jawa Tengah | 22704,50 | 7768,15 | 275,82 | 275,82 | Centroid 2 |
| 15 | Di Yogyakarta | 21008,50 | 6072,15 | 1971,82 | 1971,82 | Centroid 2 |
| 16 | Jawa Timur | 23481,50 | 8545,15 | 501,18 | 501,18 | Centroid 2 |
| 17 | Bali | 26108,50 | 11172,15 | 3128,18 | 3128,18 | Centroid 2 |
| 18 | Nusa Tenggara Barat | 21808,50 | 6872,15 | 1171,82 | 1171,82 | Centroid 2 |
| 19 | Nusa Tenggara Timur | 1048,50 | 15984,85 | 24028,82 | 1048,50 | Centroid 0 |
| 20 | Kalimantan Barat | 6716,50 | 8219,85 | 16263,82 | 6716,50 | Centroid 0 |
| 21 | Kalimantan Selatan | 12335,50 | 2600,85 | 10644,82 | 2600,85 | Centroid 1 |
| 22 | Kalimantan Tengah | 6918,50 | 21854,85 | 29898,82 | 6918,50 | Centroid 0 |
| 23 | Kalimantan Timur | 10870,50 | 4065,85 | 12109,82 | 4065,85 | Centroid 1 |
| 24 | Kalimantan Utara | 808,50 | 14127,85 | 22171,82 | 808,50 | Centroid 0 |
| 25 | Gorontalo | 13851,50 | 1084,85 | 9128,82 | 1084,85 | Centroid 1 |
| 26 | Sulawesi Selatan | 25716,50 | 10780,15 | 2736,18 | 2736,18 | Centroid 2 |
| 27 | Sulawesi Tenggara | 16085,50 | 1149,15 | 6894,82 | 1149,15 | Centroid 1 |
| 28 | Sulawesi Tengah | 22766,50 | 7830,15 | 213,82 | 213,82 | Centroid 2 |
| 29 | Sulawesi Utara | 17370,50 | 2434,15 | 5609,82 | 2434,15 | Centroid 1 |
| 30 | Sulawesi Barat | 24916,50 | 9980,15 | 1936,18 | 1936,18 | Centroid 2 |
| 31 | Maluku | 8187,50 | 23123,85 | 31167,82 | 8187,50 | C0 |
| 32 | Maluku Utara | 7024,50 | 7911,85 | 15955,82 | 7024,50 | C0 |
| 33 | Papua | 5337,50 | 20273,85 | 28317,82 | 5337,50 | C0 |
| 34 | Papua Barat | 3774,50 | 11161,85 | 19205,82 | 3774,50 | C0 |

Setelah melakukan iterasi ke-3 dengan cara yang sama, akhirnya peneliti memperoleh hasil bahwa pengelompokan data pada iterasi ke-3 sama dengan pengelompokan data pada iterasi pertama atau sebelumnya, dengan demikian proses iterasi dihentikan dengan hasil pengelompokan data harga cabai per cluster data sebagai berikut:

Cluster 0 : 10 Provinsi dengan tingkat harga cabai rawit tertinggi

Cluster 2 : 12 Provinsi dengan tingkat harga cabai rawit rendah

Cluster 1 : 12 Provinsi dengan tingkat harga cabai rawit sedang

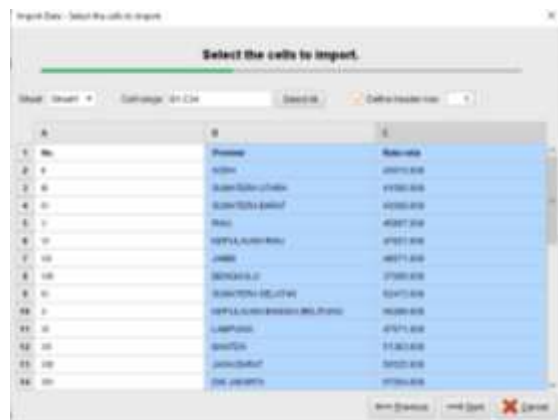
3.2 Implementasi Algoritma K-Means dengan *tool* RapidMiner

Dataset yang telah disiapkan diimpor ke dalam RapidMiner,



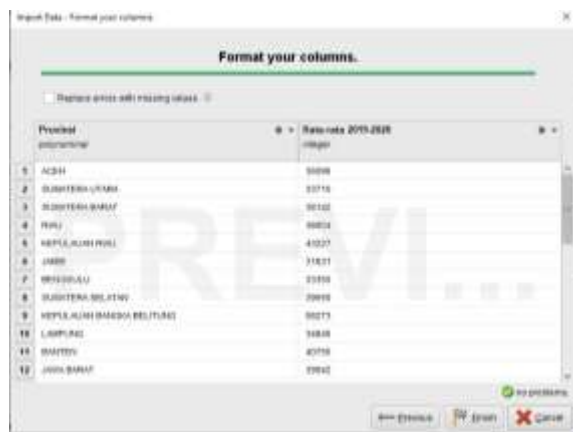
Gambar 2. Tampilan window RapidMiner import data

Kemudian lakukan seleksi data sheet Excel atau CSV.



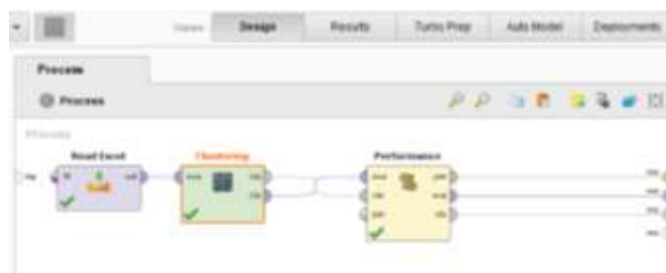
Gambar 3. Tampilan seleksi data sheet excel/csv

Jika tidak ada masalah pada data yang diimport, dapat dilanjutkan ke proses clustering data. Untuk tipe data disesuaikan dengan data yang diimport, peneliti menggunakan tipe data integer.



Gambar 4. Tampilan format data pada kolom dan tipe data

Setelah masuk ke halaman proses, drag atau masukkan Algoritma K-Means dan hubungkan ke Clustering dan res.



Gambar 5. Proses Clustering menggunakan Algoritma K-Means

Masukkan nilai K pada sebelah kanan halaman, pada penelitian ini nilai K adalah 3 karena peneliti mengelompokkan menjadi 3 cluster, kemudian masukkan nilai pada kolom max optimization steps, artinya akan dilakukan pengelompokkan data cluster hingga nilai centroid yang dihasilkan tetap tidak berubah lagi.

Cluster Model

```
Cluster 0: 10 items
Cluster 1: 12 items
Cluster 2: 12 items
Total number of items: 34
```

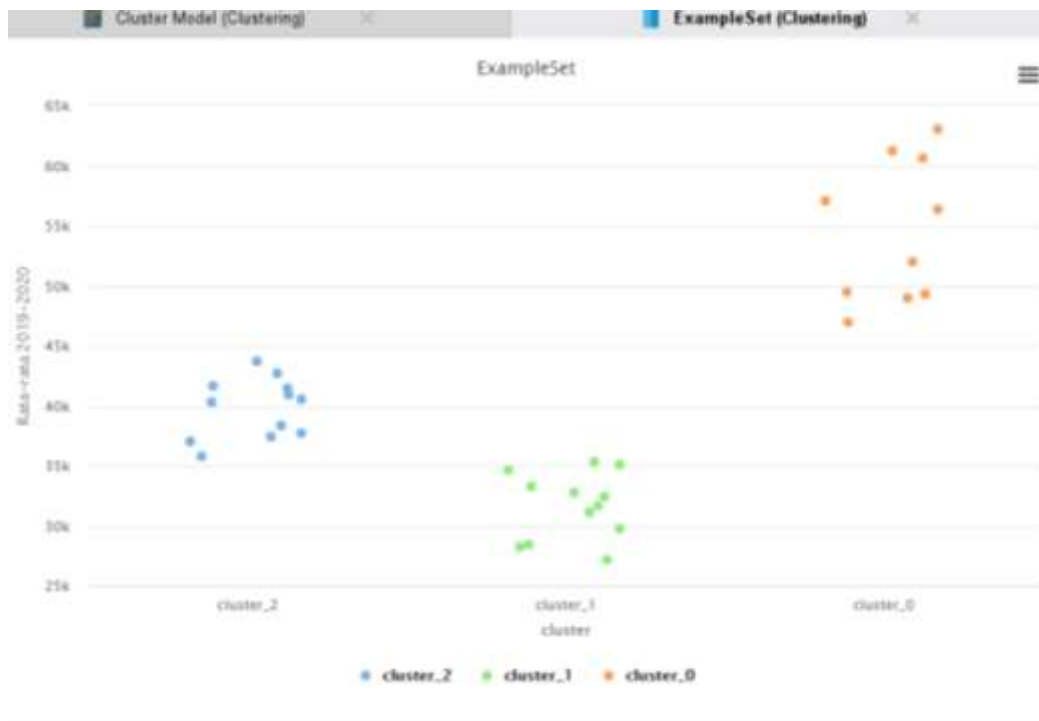
Gambar 6. Cluster Model Iterasi Pertama

Gambar 6 merupakan cluster model dari iterasi pertama. Selanjutnya peneliti melakukan lagi proses iterasi ke-2 dan hasil yang diperoleh adalah pada iterasi ke-2 menghasilkan nilai centroid pada cluster 0,1 dan 2 sama, maka proses clustering dihentikan pada iterasi ke-3.

Berikut merupakan hasil clustering dari Algoritma K-Means data pengelompokan harga cabai berdasarkan provinsi di Indonesia.

| Row No. | Provinsi | cluster | Rata-rata 20... |
|---------|--------------|-----------|-----------------|
| 1 | ACEH | cluster_2 | 30096 |
| 2 | SUMATERA... | cluster_1 | 33716 |
| 3 | SUMATERA B. | cluster_2 | 26142 |
| 4 | RIAU | cluster_2 | 38804 |
| 5 | KEPULAUAN... | cluster_2 | 43227 |
| 6 | JAWA | cluster_1 | 31831 |
| 7 | BENGKULU | cluster_1 | 33350 |
| 8 | SUMATERA S. | cluster_2 | 30650 |
| 9 | KEPULAUAN... | cluster_0 | 58273 |
| 10 | LAMPUNG | cluster_1 | 34846 |
| 11 | BANYAS | cluster_2 | 40750 |
| 12 | JAWA BARAT | cluster_2 | 39942 |

Gambar 7. Hasil clustering pengelompokan harga cabai berdasarkan provinsi



Gambar 8. Visualisasi Data Clustering

Pada gambar di atas merupakan tampilan akhir dari aplikasi K-Means dalam menentukan pengelompokan data clustering. Dengan hasil kenaikan harga cabai tertinggi pada cluster 0 merupakan kelompok provinsi dengan tingkat kenaikan harga tertinggi yaitu berjumlah 10 provinsi, cluster 1 kelompok provinsi dengan tingkat kenaikan harga terendah dan kelompok provinsi dengan tingkat kenaikan harga sedang yaitu cluster 2. Hasil data tersebut dapat dilihat pada tabel berikut :

Tabel 8. Hasil Data Clustering Kenaikan harga Cabai Rawit berdasarkan aplikasi Rapid Miner

| No | Provinsi | | |
|----|---|--|--------------------------------------|
| | Cluster 0 (Kenaikan harga tertinggi) | Cluster 1 (Kenaikan harga terendah) | Cluster 2 (Kenaikan harga sedang) |
| 1 | Kepulauan Babel | Sumatera Utara | Aceh |
| 2 | DKI Jakarta | Jambi | Sumatera Barat |
| 3 | NTT | Bengkulu | Riau |
| 4 | Kalimantan Barat | Lampung | Kepulauan Riau |
| 5 | Kalimantan Tengah | Jawa Tengah | Sumatera Selatan |
| 6 | Kalimantan Utara | DI Yogyakarta | Banten |
| 7 | Maluku | Jawa Timur | Jabar |
| 8 | Maluku Utara | Bali | Kalimantan Selatan |
| 9 | Papua | NTB | Kalimantan Timur |
| 10 | Papua Barat | Sulawesi Selatan | Gorontalo |
| 11 | | Sulawesi Tengah | Sulawesi Tenggara |
| 12 | | Sulawesi Barat | Sulawesi Utara |

4. KESIMPULAN

Berdasarkan penelitian yang dilakukan, maka dapat disimpulkan hasil pengujian pengelompokan data kenaikan harga cabai pada pertengahan tahun bulan Juli 2020 hingga Juli 2021 di seluruh provinsi baik dengan perhitungan manual yang dilakukan melalui proses 3 iterasi maupun menggunakan *tool* RapidMiner mampu menyelesaikan persoalan tersebut dengan menghasilkan pengelompokan data yang sama dimana terdiri dari 3 Cluster tingkatan harga cabai rawit ditandai dengan 3 jenis warna yang mewakili 3 cluster harga; tertinggi, sedang dan terendah. Sehingga menjadi bahan evaluasi atau dasar pertimbangan pihak-pihak terkait membuat kebijakan dalam hal menjamin ketersediaan, mengurangi disparitas harga, dan menjaga kestabilan harga baik pada saat harga tinggi maupun pada saat harga jatuh. Kedepan, dapat dilakukan analisis pengelompokan data dengan alat analisis dan algoritma yang berbeda agar mengetahui cara kerja serta hasil keputusan yang bervariasi.

DAFTAR PUSTAKA

- [1] H. Arizka, I. Hasan, and I. Rosada, "Analisis Faktor-Faktor Yang Mempengaruhi Fluktuasi Harga Cabai Rawit Di Pasar Barandasi, Kabupaten Maros," *Wiratani*, vol. 1, no. 2, pp. 116–125, 2018.
- [2] A. K. Wardhani, "K-Means Algorithm Implementation for Clustering of Patients Disease in Kaje Clinic of Pekalongan," *J. Transform.*, vol. 14, no. 1, p. 30, 2016, doi: 10.26623/transformatika.v14i1.387.
- [3] Z. Nabila, "Analisis Data Mining Untuk Clustering Kasus Covid-19," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, pp. 100–108, 2021.
- [4] M. A. K-means, A. Novita, and H. B. Seta, "Pemetaan Pasar Tradisional Berdasarkan Harga Pangan Komoditas," no. September, pp. 320–326, 2021.
- [5] P. Alkhairi and A. P. Windarto, "Penerapan K-Means Cluster pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara," *Semin. Nas. Teknol. Komput. Sains*, pp. 762–767, 2019.
- [6] M. H. Siregar, "Data Mining Klasterisasi Penjualan Alat-Alat Bangunan Menggunakan Metode K-Means (Studi Kasus Di Toko Adi Bangunan)," *J. Teknol. Dan Open Source*, vol. 1, no. 2, pp. 83–91, 2018, doi: 10.36378/jtos.v1i2.24.
- [7] T. Siburian, M. Safii, and I. Parlina, "Penerapan Algoritma K-Means Clustering untuk Pengelompokan Harga Eceran Beras di Pasar Tradisional Berdasarkan Wilayah Kota," *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 1, p. 927, 2019, doi:

10.30645/senaris.v1i0.101.

- [8] R. Wulan Sari, A. Wanto, and A. Perdana Windarto, “KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer) IMPLEMENTASI RAPIDMINER DENGAN METODE K-MEANS (STUDY KASUS: IMUNISASI CAMPAK PADA BALITA BERDASARKAN PROVINSI),” vol. 2, no. 1, pp. 224–230, 2018, [Online]. Available: <http://ejurnal.stmik-budidarma.ac.id/index.php/komik>.
- [9] D. G. Ramadhan, I. Prihatini, and F. Liantoni, “Analisis Clustering Pengelompokan Penjualan Paket Data Menggunakan Metode K-Means,” *Ultim. J. Tek. Inform.*, vol. 13, no. 1, pp. 33–38, 2021, doi: 10.31937/ti.v13i1.1981.
- [10] Y. Prayoga, H. S. Tambunan, and I. Parlina, “Penerapan Clustering Pada Laju Inflasi Kota Di Indonesia Dengan Algoritma K-Means,” *BRAHMANA J. Penerapan Kecerdasan Buatan*, vol. 1, no. 1, pp. 24–30, 2019, doi: 10.30645/brahmana.v1i1.4.