

Klasterisasi Kualitas Biji Kopi Berdasarkan Taraf Penyusutan Menggunakan Metode *K-Harmonic Means* dengan Validasi *Silhouette Index* dan *C-index*

Arwin Putra¹, Dahlan Abdullah², Muhammad Daud^{3*}

^{1,2,3}Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Malikussaleh

*e-mail: arwinputragayo25@gmail.com¹, dahlan@unimal.ac.id², mdaud@unimal.ac.id³

Received:	Revised:	Accepted:	Available online:
21.08.2024	11.09.2024	19.10.2024	30.10.2024

Abstrak: Kopi merupakan salah satu hasil perkebunan masyarakat Indonesia yang memiliki nilai jual yang tinggi. di Indonesia ada beberapa daerah penghasil kopi, salah satunya adalah Kabupaten Bener Meriah dengan luas lahan 48.95 ribu ha dan jumlah produksi 4.75 ribu ton. Setiap desa nya memiliki jumlah produktifitas kopi asalan yang berbeda-beda, walaupun dengan luas lahan yang sama, hal ini terjadi karena adanya proses penyusutan. Oleh karena itu perlu adanya pengelompokan desa dengan kualitas biji kopi berdasarkan taraf penyusutan. Penelitian ini bertujuan untuk mengelompokkan desa-desa dan mengetahui *cluster* kualitas biji kopi berdasarkan proses pengolahan. Metode yang digunakan dalam pengelompokan (*Clustering*) ini menggunakan metode *K-Harmonic Means* yang merupakan pengembangan dari metode *K-Means*. Jumlah data yang digunakan dalam penelitian ini yaitu sebanyak 34 desa dengan 6 atribut yaitu nama desa, luas lahan, jumlah gelondong, gabah, labu dan asalan . pengujian klaster dilakukan dengan jumlah klaster $k = 2, 3$ dan 4 serta menguji klaster yang dihasilkan dengan menggunakan metode validasi *silhouette index* dan *C-Index*. Berdasarkan hasil analisis penelitian ini didapatkan klastering yang terbaik pada $k = 2$ dengan metode *Silhouette Index* dengan nilai $SI = 0,59822$ sedangkan *c-index* diperoleh klaster yang terbaik pada $k = 4$ dengan nilai $CI = 0,03436$. Adapun hasil profilisasi $k = 2$ yaitu 9 desa dengan kualitas kopi sangat baik dan 25 desa dengan kualitas kopi baik. dan profilisasi $k = 4$ yaitu 5 desa dengan kualitas kopi sangat baik, 13 desa dengan kualitas kopi baik, 10 desa dengan kualitas kopi cukup dan 6 desa dengan kualitas kurang. Dari kedua metode ini yang paling optimal yaitu menggunakan 4 klaster pada *c-index* yang memiliki nilai rasio Sw dan Sb terkecil yaitu 0,0734.

Kata kunci: kopi, k-harmonic means, pengklasteran, silhouette index, c-index

Abstract: Coffee is one of the plantation products of Indonesian people which has a high selling value. In Indonesia there are several coffee producing areas, one of which is Bener Meriah Regency with a land area of 48.95 thousand ha and a production volume of 4.75 thousand tons. Each village has a different amount of asalan coffee productivity, even with the same land area, this occurs because of the shrinkage process. Therefore, it is necessary to group villages according to the quality of coffee beans based on the level of shrinkage. This research aims to group villages and determine clusters of coffee bean quality based on the processing process. The method used in clustering uses the *K-Harmonic Means* method which is a development of the *K-Means* method. The amount of data used in this research was 34 villages with 6 attributes, namely village name, land area, number of spindles, grain, pumpkins and random. Cluster testing was carried out with the number of clusters $k = 2, 3$ and 4 and tested the resulting clusters using the *Silhouette Index* and *C-Index* validation methods. Based on the results of this research analysis, the best clustering was obtained $k = 2$ using the *Silhouette Index* method with a SI value = 0.59822, while the *c-index* obtained the best cluster at $k = 4$ with a CI value = 0.03436. The results of the $k = 2$ profile are 9 villages with very good coffee quality and 25 villages with good coffee quality. and profile $k = 4$, namely 5 villages with very good coffee quality, 13 villages with good coffee quality, 10 villages with sufficient coffee quality and 6 villages with poor quality. Of these two methods, the most optimal is using 4 clusters in the *c-index* which has the smallest Sw and Sb ratio value, namely 0.0734.

Keywords: coffee, k-harmonic means, clustering, silhouette index, c-index

1. PENDAHULUAN

Kopi merupakan salah satu hasil perkebunan masyarakat Indonesia yang memiliki nilai jual yang tinggi [1]. Salah satu ekspor yang penting dalam perdagangan dunia adalah kopi, dan Indonesia merupakan Negara keempat terbesar pengeksport kopi [2]. Indonesia ada beberapa daerah penghasil kopi Arabika, salah satunya adalah Kabupaten Bener Meriah, Provinsi Aceh dengan luas lahan 48.95 ribu hektar dan jumlah produksi 4.75 ribu ton [3]. Komoditas unggulan pertanian di Kabupaten Bener Meriah adalah dari sektor perkebunan tanaman kopi yaitu kopi jenis arabika. Kualitas produktifitas kopi kabupaten ini juga tetap di tingkatkan dan di berikan perawatan secara merata di setiap daerah.

Setiap daerah di Bener Meriah memiliki jumlah produktifitas kopi asalan yang berbeda-beda, walaupun dengan luas lahan yang sama di setiap daerah/desa. Hal ini terjadi karena adanya penyusutan dari proses pengolahan kopi yang dimulai dari pemetikan buah *cherry* (gelondong), buah *cherry* di *pulping* akan menghasilkan kopi Gabah, kopi gabah diproses *hulling* akan menghasilkan

kopi labu, selanjutnya kopi labu diproses menjadi kopi asalan. Dari proses tersebut terjadi penyusutan yang berbeda di setiap daerah dimana proses penyusutan ini biasanya di sebut dengan rendemen. Oleh karena itu dari proses penyusutan tersebut perlu adanya pengelompokan daerah/desa dengan kualitas kopi berdasarkan taraf penyusutan.

Salah satu metode yang digunakan dalam pengelompokan /*clustering* ini adalah metode yang terdapat dalam data mining yaitu *K-Harmonic Means*. *K-Harmonic Means* merupakan pengembangan dari metode *K-Means*. Perbedaan *K-Harmonic Means* dengan *K-means* adalah penggunaan nilai titik pusat klaster, dimana metode *K-Harmonic Means* dalam menentukan keberadaan tiap titik datanya berdasarkan dari fungsi keanggotaan, sehingga menyebabkan metode ini bersifat tidak sensitif dalam inisialisasi awal titik pusat klaster dan secara signifikan meningkatkan kualitas klasterisasi dibandingkan dengan *k-means* [4]. Dalam menganalisis klaster yang terbentuk perlu adanya validasi hasil pengklasteran dengan menggunakan dua validasi yaitu validasi *silhouette index* dan *C-Index*. diharapkan dengan adanya hasil dari penelitian ini dapat menghasilkan pengetahuan informasi mengenai kelompok atau klaster kualitas biji kopi dimulai dari proses *cherry* (gelondong), gabah, labu dan asalan. Sehingga diketahui daerah yang termasuk dalam klaster kualitas kopi. Berdasarkan informasi klaster tersebut dapat ditentukan dan di arahkan prioritas perawatan kopi disetiap daerah.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data mining merupakan proses dengan kecerdasan buatan, teknik statistik, matematika, dan *machine learning* untuk mengidentifikasi dan mengekstraksi suatu informasi [5]. Tujuan data mining adalah untuk menemukan, menggali, atau menambang pengetahuan dari data atau informasi yang kita miliki [6]. Metode seperti *K-Means* dan DBSCAN masuk dalam kategori pengelompokan *partitioning* [7].

2.1.1. Algoritma KHM

Algoritma *K-means* sangat membantu dalam mensegmentasi suatu populasi klien pemulihan sensitif ke dalam subkelompok yang lebih homogen [8]. *K-Means* tidak menjamin hasil klasterisasi yang unik karena metode ini dapat menghasilkan hasil klaster yang berbeda tergantung dari posisi inisialisasi klaster awal [9]. Tujuan pengembangan metode KHM adalah untuk menangani masalah utama dalam *K-Means* yang hasil *clustering* nya sangat sensitif dengan inisialisasi data yang dijadikan sebagai *centroid* awal. Hasil yang sering berbeda (lokal optima) dari proses *clustering* nya (pada set data yang sama) disebabkan oleh inisialisasi *centroid* yang berbeda [10].

2.2. Analisis Silhouette

Analisis *Silhouette* mengacu pada metode interpretasi dan validasi konsistensi dalam kelompok data. Menghitung koefisien *silhouette* yang didefinisikan sebagai rata-rata $s(i)$ [11]. Semua indeks validitas klaster ini menggabungkan informasi tentang kekompakan intracluster dan isolasi antar cluster, serta faktor lain, seperti sifat geometris atau statistik data, jumlah objek data dan pengukuran ketidaksamaan atau kemiripan [12].

2.3. Analisis Silhouette

Principal Componen Analysis (PCA) adalah salah satu metode yang dikenal luas dan digunakan untuk mereduksi dimensi data multivariat, kompresi data, pengenalan pola pada jaringan saraf, dan analisis statistik [13].

2.4. Penelitian Terkait

Berikut ini beberapa penelitian terkait mengenai metode *K-Harmonic Means* serta validasi *silhouette index* dan *C-Index*, pada penelitian [14] Penerapan Metode *K – Harmonic Means* Dalam Pengelompokan Kabupaten/Kota di peroleh 2 klaster dan menggunakan parameter (p) sebesar 4. Serta nilai *Silhouette Coefficient* yang didapat sebesar 0,323 yang dikategorikan kedalam *weak structure* (struktur klaster yang lemah).

Pada penelitian Poerwanto [15] Penelitian ini dilakukan pada tahun 2021 yang bertujuan untuk mengelompokkan 45 kecamatan di kabupaten tersebut menjadi 2 sampai 5 klaster berdasarkan

produktivitasnya dalam menyediakan tanaman perkebunan menggunakan Algoritma K-Means. metode validasi kluster yaitu *Silhouette Index*, dan nilai yang didapat sebesar 0.8068 dengan 2 kluster yang optimal.

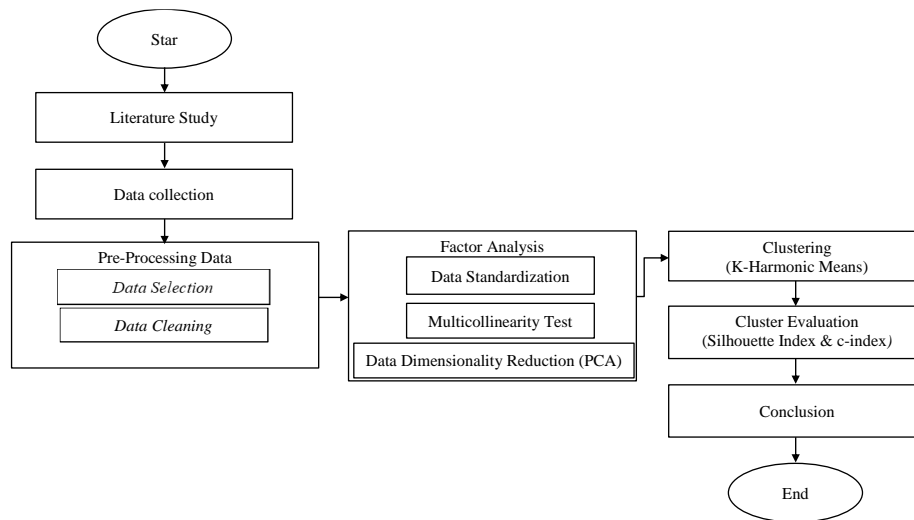
Pada penelitian selanjutnya yang dilakukan oleh H. Mahi [16] Tujuan dari penelitian ini adalah menyelidiki kinerja *Silhouette indeks* dengan menerapkan algoritma *K-Harmonic Means*. Hasil dari penelitian ini yaitu antara algoritma *K-Harmonic Means* dan *Silhouette indeks* cocok untuk citra penginderaan jauh yang menyajikan tingkat tumpang tindih antar kluster yang tidak signifikan.

Penelitian selanjutnya yang dilakukan oleh A. Assegaf [17] ". Mengklasifikasikan kesehatan bank menggunakan metode *Local Mean K-Nearest Neighbor* dan *Multi Local Means K-Harmonic Nearest Neighbor*. Selanjutnya pada penelitian T. P. Fiqar [18] Tujuan dari penelitian ini adalah mengembangkan metode *K – Harmonic Means* menggunakan *Improved Artificial Bee Colony (IABC)* dengan hasil berhasil mengoptimalkan posisi titik pusat kluster yang mengarahkan hasil kluster yang memiliki solusi global.

3. METODE

3.1 Penelitian Terkait

Tahap penelitian ini merupakan alur penelitian yang akan dilakukan dalam melakukan proses *clustering* kualitas biji kopi berdasarkan taraf penyusutan menggunakan metode *K-harmonic means*. Secara umum tahapan penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

3.2 Studi Literatur

Studi literatur yang dilakukan dalam penelitian ini adalah menguraikan penelitian terkait dan teori-teori yang menjadi pendukung penelitian dari berbagai sumber seperti konsep *data mining*, metode *K-Means*, *K-Harmonic Means*, metode validasi kluster dan lain-lain.

3.3 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data primer yaitu penjualan kopi dari kolektor kepada koperasi permata gayo tahun 2020 s/d 2021 dari 34 desa seperti ditunjukkan dalam Tabel 1. Adapun metode pengumpulan data yang digunakan dalam penelitian ini adalah observasi dan wawancara yang dilakukan langsung dengan pihak koperasi permata gayo untuk memperoleh data dan informasi tentang penjualan kopi. Data yang didapat dalam bentuk file *excel* sehingga berekstensi *.xlsx*.

Tabel 1. Data mentah penjualan kopi dari kolektor

Desa (FB)	Kolektor	Gelondong		Gabah		Huller	Gabah/labu		Asalan	Tanggal Masuk Asalan Ke D10
		Tanggal	Bambu	Jumlah	Bambu	Kg	Tanggal	Kg		

		Rekap		Goni							
Temas Mumanang	11	M. Saleh	05/01/2020	16.072	76	6.107	D-06	4.886	11/01/2020	2.893	16/01/2020
Temas Mumanang	11	M. Saleh	10/01/2020	13.911	70	5.564	D-06	4.507	16/01/2020	2.643	16/01/2020
Tanjung Sari	15	Zukardi	05/01/2020	14.317	73	5.870	D-10	4.696	11/01/2020	2.577	16/01/2020
Tanjung Sari	15	Zukardi	10/01/2020	14.870	74	5.948	D-10	4.877	16/01/2020	2.974	16/01/2020
Suku Rawe	13	Arifuddin	05/01/2020	13.889	73	5.833	D-06	4.667	11/01/2020	2.500	16/01/2020
Suku Rawe	13	Arifuddin	10/01/2020	12.585	63	5.034	D-06	4.178	16/01/2020	2.517	17/01/2020
Bintang Bener Mutiara	18	Lukman	08/01/2020	10.229	50	3.989	D-11	3.231	14/01/2020	2.148	17/01/2020
...
...
...
Oneng	21	Ridwan Ahmad	24/01/2020	3.944	19	1.538	D-06	1.231	29/01/2020	710	29/01/2020
Rembele Timur	38	Ishak	24/01/2020	16.305	84	6.685	D-10	5.348	29/01/2020	3.424	29/01/2020
Vitran TM	41	Saukani	24/01/2020	6.753	34	2.701	D-06	2.215	29/01/2020	1.283	29/01/2020

3.4 Preprocessing Data

Pada tahap ini memproses data mentah yang telah diperoleh, sebelum data diproses menggunakan metode *K-Harmonic Means*, terlebih dahulu harus melewati tahap ini. *Preprocessing* data yang dilakukan pada data penelitian ini adalah:

- a. Data seleksi dimana mengurangi atau menghilangkan beberapa variabel yang tidak digunakan pada penelitian ini seperti variabel kolektor, tanggal rekap, huller, jumlah goni, tanggal, dan tanggal masuk. Hasil dari data seleksi dapat dilihat pada Tabel 2 berikut.

Tabel 2. Hasil proses data seleksi

Desa (FB)	Gelondong (Bambu)	Gabah (Bambu)	Gabah/Labu (Kg)	Asalan (Kg)
Temas Mumanang	16.072	6.107	4.886	2.893
Temas Mumanang	13.911	5.564	4.507	2.643
Tanjung Sari	14.317	5.870	4.696	2.577
Tanjung Sari	14.870	5.948	4.877	2.974
Suku Rawe	13.889	5.833	4.667	2.500
Suku Rawe	12.585	5.034	4.178	2.517
Bintang Bener Mutiara	10.229	3.989	3.231	2.148
...
...
Rembele Timur	16.305	6.685	5.348	3.424
Vitran TM	6.753	2.701	2.215	1.283

- b. Data *Integration* yaitu menggabungkan data gelondong, gabah, labu dan asalan berdasarkan kolektor dan desa. Hasil dari data yang sudah dilakukan *Preprocessing* dapat dilihat pada Tabel 3 berikut.

Tabel 3. Data yang sudah di *Preprocessing*

No.	Desa	Luas Lahan (Ha)	Gelondong (Bambu)	Gabah (Bambu)	Labu (Kg)	Asalan (Kg)
1	Temas Mumanang	66,5	391.648	158.319	128.451	80.648
2	Suku Rawe	111	691.536	283.309	230.152	144.284
3	Tanjung Sari	80,75	561.865	230.131	186.796	118.248
4	Bener Lukup II	92	505.265	205.195	166.364	103.585
5	Bintang Bener Mutiara	171	748.833	304.051	246.620	157.201
6	Oneng	103	567.313	230.259	187.193	118.096
7	Panji 1	21,5	147.550	60.032	48.799	30.200
---	---	---	---	---	---	---
---	---	---	---	---	---	---

27	Teleden	31,8	211.000	84.007	68.238	44.194
28	Sumber Rejo	20	242.416	99.093	80.568	49.989
29	Wonosari	44,055	406.303	166.705	133.114	91.445
30	Tansaran Bidin	32,5	343.671	144.067	115.376	78.664
31	Kekuyang	34,26	395.230	163.155	131.796	89.618
32	Bakongan	32,75	237.244	94.787	75.112	49.453
33	Timur Jaya	36	205.818	83.795	66.962	43.047
34	Reje Guru	22,24	171.151	71.527	57.487	37.102

3.5 Analisis Faktor

Setelah data selesai dilakukan *preprocessing*, selanjutnya data akan dilakukan reduksi dimensi guna untuk mempermudah proses klastering menggunakan metode *K-Harmonic means*. Data akan direduksi dari 5 variabel menjadi 2 faktor utama menggunakan metode *Principal Component Analysis (PCA)*. Sebelum dilakukan reduksi data, terlebih dahulu harus melakukan standarisasi data dengan metode *z-score* menggunakan persamaan berikut.

$$Z = \frac{x - \mu}{\sigma} \tag{1}$$

dimana:

Z = nilai *standart score*

x = data observasi

μ = rata-rata variabel

σ = rata-rata variabel

Standarisasi ini bertujuan agar interval data menjadi lebih proporsional dengan menyamakan rentang nilai pada semua data. Setelah ini melakukan uji multikolinieritas menggunakan metode KMO untuk melihat ketergantungan variabel satu dengan variabel lainnya

3.6 K-Harmonic Means

Melakukan proses klaster pada data penjualan kopi menggunakan metode *K-harmonic means*. Variabel-variabel yang digunakan dalam penelitian ini adalah desa, gelondong, gabah, labu dan asalan. *K-Harmonic Means* yang diolah menggunakan menggunakan *Software* yaitu Aplikasi RStudio. Data yang digunakan dalam pengujian ini sebanyak 34 desa dengan variabel atau 2 komponen utama. Percobaan Pengujian dilakukan dengan menggunakan variasi dengan nilai $k=2, 3$ dan 4 . Adapun langkah-langkah yang dilakukan penulis dalam mengelompokan kualitas biji kopi berdasarkan taraf penyusutan menggunakan metode *K-Harmonic Means* adalah sebagai berikut:

1. Menentukan jumlah data yang akan di *cluster*, dimana sampel data yang akan digunakan dalam proses *clustering* adalah data penjualan kopi dari Kolektor kepada Koperasi Permata Gayo.
2. Menentukan pusat klaster (*centroid*) awal, dari 34 data. Jumlah klaster yang digunakan yaitu $k=2, 3$ dan 4 .
3. Melakukan analisis klaster dengan algoritma *k-harmonic means*:
 - a. Membangkitkan inisialisasi posisi titik pusat klaster C (k) *centroid* awal sebanyak k *centroid*.
 - b. Selanjutnya untuk setiap data x_i , hitung nilai keanggotaan $m(c_l|x_i)$ untuk setiap titik pusat klaster c_l berdasarkan persamaan berikut.

$$KHM(X, C) = \sum_{i=1}^N \frac{K}{\sum_{j=1}^K \frac{1}{\|x_i - c_j\|^p}} \tag{2}$$

dimana:

k = klaster

p = input parameter. Nilai p biasanya ≥ 2 .

X_i = data ke- i

c_j = titik pusat klaster ke- j

$KHM(X, C)$ = nilai fungsi keanggotaan KHM.

- c. Untuk setiap data x_i , hitung nilai bobot $m(x_i)$ dengan persamaan berikut ini

$$m(c_l|x_i) = \frac{\|x_i - c_l\|^{-p-2}}{\sum_{i=1}^k \|x_i - c_l\|^{-p-2}} \tag{3}$$

- d. Untuk setiap titik pusat x_i , ulang kembali perhitungan untuk posisi pusat kluster dari semua data berdasarkan nilai keanggotaan dan bobot yang dimiliki tiap data. Penentuan posisi titik pusat ini berdasarkan persamaan berikut

$$c_l = \frac{\sum_{i=1}^N m(c_l|x_i) \cdot w(x_i) \cdot x_i}{\sum_{i=1}^N m(c_l|x_i) \cdot w(x_i)} \quad (4)$$

dimana:

- c_j = nilai titik pusat kluster baru
- $m(c_j|x_i)$ = nilai keanggotaan data x terhadap titik pusat kluster c
- $w(X_i)$ = nilai bobot x_i

- e. Ulangi perhitungan nilai fungsi tujuan sampai *update* centroid mendapatkan nilai fungsi tujuan yang tidak terdapat perubahan atau kurang dari ambang batas.
- f. Tetapkan keanggotaan data x_i pada suatu kluster dengan titik pusat kluster c_j sesuai dengan nilai keanggotaan x_i terhadap c_j .
4. Melakukan validasi hasil pengklasteran
- a. Menghitung nilai *Silhouette Index* dan *C-index* pada masing-masing kluster yang terbentuk yaitu dari nilai $k=2, 3$ dan 4 .
 - b. Membandingkan nilai *Silhouette Index* dan *C-index* dari $k = 2, 3$ dan 4 pengklasteran.
 - c. Hasil kluster yang terbaik ditunjukkan dengan nilai maksimum pada *Silhouette Index* dan minimum pada *C-index*.
5. Melakukan interpretasi terhadap hasil klustering terhadap kualitas kopi berdasarkan taraf penyusutan dari 34 desa.

3.7 Evaluasi

Pada tahap ini, hasil kluster yang sudah terbentuk akan diuji dan divalidasi hasil pengklasteran dengan menggunakan dua validasi dengan pendekatan kriteria internal yaitu validasi *silhouette index* dan *C-Index* yang dihitung menggunakan persamaan berikut.

$$S(i) = \frac{b(i)-a(i)}{\max(a(i),b(i))}, b(i) = \min d(i, c), \text{ dan } a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (5)$$

$b(i)$ = nilai minimum dari jarak rata-rata objek i dengan semua objek pada kluster lain C
 $a(i)$ = rata-rata jarak objek ke- i dengan semua objek yang berada di dalam satu kluster A

$$C - Index = \frac{S_w - S_{min}}{S_{max} - S_{min}}, S_{max} \neq S_{min} \quad (6)$$

Dimana:

- S_{min} =Jumlah jarak nilai minimum dalam kluster dan antar kluster
- S_{max} =Hitung jumlah jarak nilai maksimum dalam kluster dan antar kluster

4. HASIL DAN PEMBAHASAN

Penelitian ini merupakan Analisa evaluasi validasi kluster menggunakan metode validasi *silhouette index* dan *c-index*, pada penerapan metode *K-Harmonic Means* menggunakan bahasa pemrograman R dibantu menggunakan *tool* RStudio. Data yang akan digunakan adalah data penjualan kopi dari 34 desa yang ada di Kabupaten Bener Meriah. Metode *K-Harmonic Means* diterapkan untuk mengelompokkan setiap desa berdasarkan data penjualan kopi tersebut langkah awal yang dilakukan adalah menentukan jumlah kluster yang ingin dibentuk, dalam penelitian ini nilai k yang digunakan yaitu $k= 2, 3$ dan 4 . Selanjutnya menentukan *centroid* awal, menghitung jarak antara objek ke setiap *centroid* yang sudah ditentukan dan menghasilkan kluster (iterasi sampai anggota atau *centroid* tidak berubah). Selanjutnya akan dilakukan validasi kluster menggunakan dua metode sekaligus yaitu metode validasi *silhouette index* dan *c-index* dan membandingkan nilai dari masing-masing *indeks* validasi tersebut. hasil standarisasi data dapat dilihat pada Tabel 4.

4.1 Standarisasi Data

Dari dataset penelitian ini, selanjutnya adalah melakukan standarisasi data dari variabel luas lahan, gelondong, gabah, labu dan asalan, yang dilakukan dengan menggunakan rumus standarisasi data yaitu *z-score*. Standarisasi data menggunakan *z-score* yang dilakukan dengan *software* R-Studio dengan fungsi *scale()*.

Tabel 4. Hasil Standarisasi data

No.	Desa	Luas Lahan (Ha)	Gelondong (Bambu)	Gabah (Bambu)	Labu (Kg)	Asalan (Kg)
1	Temas Mumanang	0,221	-0,019	-0,039	-0,037	-0,060
2	Suku Rawe	1,527	1,719	1,736	1,738	1,719
3	Tanjung Sari	0,639	0,967	0,981	0,981	0,991
4	Bener Lukup II	0,969	0,639	0,627	0,625	0,581
5	Bintang Bener Mutiara	3,287	2,051	2,031	2,025	2,080
6	Oneng	1,292	0,999	0,983	0,988	0,987
7	Panji I	-1,099	-1,434	-1,435	-1,427	-1,470
---	---	---	---	---	---	---
---	---	---	---	---	---	---
27	Teleden	-0,797	-1,066	-1,094	-1,088	-1,079
28	Sumber Rejo	-1,143	-0,884	-0,880	-0,872	-0,917
29	Wonosari	-0,437	0,065	0,080	0,045	0,242
30	Tansaran Bidin	-0,776	-0,298	-0,241	-0,265	-0,115
31	Kekuyang	-0,725	0,001	0,030	0,022	0,191
32	Bakongan	-0,769	-0,914	-0,941	-0,968	-0,932
33	Timur Jaya	-0,674	-1,096	-1,097	-1,110	-1,111
34	Reje Guru	-1,077	-1,297	-1,271	-1,275	-1,277

4.2 Reduksi Dimensi

Proses ini bertujuan untuk mereduksi variabel penelitian ini menjadi lebih sedikit atau merubah dimensi *dataset* menjadi 2 dimensi dengan membentuk variabel baru yang disebut dengan *principal component* (PC) dan tanpa harus mengurangi informasi yang termuat dalam data asli, reduksi dimensi ini dilakukan karena nantinya hasil klaster yang diperoleh akan ditampilkan dalam plot 2 dimens, oleh karena itu perlu dilakukan reduksi variabel penelitian dari 5 variabel menjadi 2 variabel. Reduksi dimensi ini dilakukan menggunakan metode *Principal Component Analysis* (PCA), dengan perhitungan *eigen value* dan *eigen vector* yang diolah menggunakan *software r-studio* dengan *library(factoextra)* dan cukup menggunakan fungsi *prcomp*. Hasil PCA dapat dilihat pada Tabel 5 berikut.

Tabel 5. Hasil PCA

No.	Desa	PC1	PC2
1	Temas Mumanang	-0,02156	-0,234
2	Suku Rawe	-3,77788	0,047341
3	Tanjung Sari	-2,04875	0,233472
4	Bener Lukup II	-1,52717	-0,36893
5	Bintang Bener Mutiara	-5,08928	-1,28793
6	Oneng	-2,33645	-0,35321
7	Panji I	3,078728	-0,19891
8	Batin Jaya	-0,84377	-0,77758
9	Penosan Jaya	0,113508	-0,50509
10	Wih Resap	-0,07201	-0,14572
11	Musara Ate	-4,02294	0,223848
12	Hakim Wih Hilang	1,676516	0,167227
13	Bener Pepanyi	1,151294	0,262513
14	Ayu Ara	1,307209	0,35313
15	Delong Asli	0,409352	0,128393
16	Ujung Gele	1,086372	-0,38512
17	Belang Tampu	2,737029	-0,57491
18	Waq Pondok Sayur	-1,05453	0,447113
19	Rembele Timur	1,484039	-0,12361
20	Jadi Sepakat	0,287525	0,431109
21	Tingkem	1,590672	0,032163
22	Vitran Temas Mumanang	0,628081	-0,4636
23	Petukel	-0,64556	0,0676
24	Bukit Nosar	1,130392	0,335832

25	Selamat Rejo	-3,4658	1,081133
26	Sedie Jadi	-4,17235	0,229307
27	Teleden	2,298915	-0,17439
28	Sumber Rejo	2,091001	0,300667
29	Wonosari	-0,0148	0,490571
30	Tansaran Bidin	0,740622	0,518107
31	Kekuyang	0,190419	0,712822
32	Bakongan	2,027132	-0,08088
33	Timur Jaya	2,287438	-0,30464
34	Reje Guru	2,776603	-0,08382

Hasil pengelompokan variable kedalam komponen utama 1 (PC1) adalah Luas Lahan, sedangkan komponen utama 2 (PC2) terdiri dari gelondong, gabah, labu dan asalan.

4.3 Hasil Klastering dengan K-Harmonic Means

Setelah data selesai direduksi dan sudah didapatkan hasil komponen utama, maka selanjutnya adalah melakukan proses klastering terhadap data penjualan kopi dengan jumlah desa sebanyak 34 desa menggunakan metode *K-Harmonic Means*. Proses klasterisasi ini dilakukan dengan percobaan jumlah klaster (k) yang berbeda-beda yaitu dengan nilai $k=2$, $k=3$ dan $k=4$. Hal ini dilakukan untuk mengetahui jumlah klaster yang paling optimal dalam mengelompokkan desa-desa berdasarkan hasil penjualan kopi.

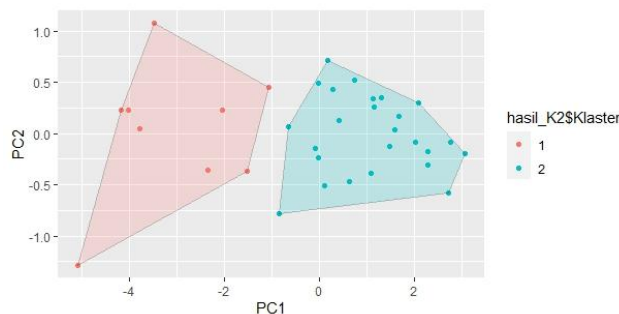
4.3.1 Hasil Klastering dengan K-Harmonic Means

Pengujian pertama dilakukan dengan jumlah klaster sebanyak 2 klaster yang diuji terhadap data penjualan kopi sebanyak 34 desa. Adapun hasil klaster akhir pada pengujian ini dapat dilihat pada Tabel 6 berikut.

Tabel 6. Hasil klaster dengan nilai $k=2$

No	Desa	PC1	PC2	Klaster
1	Suku Rawe	-3,7779	0,04734	1
2	Tanjung Sari	-2,0487	0,23347	1
3	Bener Lukup II	-1,5272	-0,3689	1
4	Bintang Bener Mutiara	-5,0893	-1,2879	1
5	Oneng	-2,3365	-0,3532	1
6	Musara Ate	-4,0229	0,22385	1
7	Waq Pondok Sayur	-1,0545	0,44711	1
---	---	---	---	---
---	---	---	---	---
30	Tansaran Bidin	0,74062	0,51811	2
31	Kekuyang	0,19042	0,71282	2
32	Bakongan	2,02713	-0,0809	2
33	Timur Jaya	2,28744	-0,3046	2
34	Reje Guru	2,7766	-0,0838	2

Hasil klaster dari tabel 6 juga dapat dilihat dengan bentuk visualisasi yang diplot menggunakan fungsi *ggplot()*, yang merupakan suatu fungsi dari *library(ggplot2)*. Visuasilasi klaster dengan nilai $k=2$ dapat dilihat pada Gambar 2 berikut.



Gambar 2. Visualisasi klaster dengan nilai $k=2$

Pada visualisasi dapat dilihat bahwa 2 kluster ini dibedakan antara dua warna yaitu merah dan biru, warna merah menunjukkan desa tersebut berada pada kluster 1 yaitu sebanyak 9 desa. Sedangkan warna biru menunjukkan desa berada pada kluster 2 yaitu sebanyak 25 desa

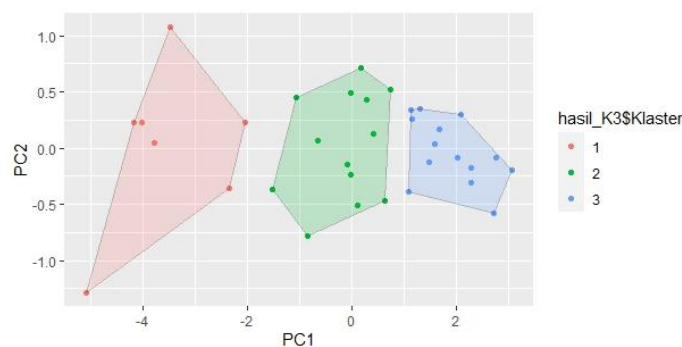
4.3.2 Pengujian Kluster k=3

Pengujian pertama dilakukan dengan jumlah kluster sebanyak 3 kluster yang diuji terhadap data penjualan kopi sebanyak 34 desa. Adapun hasil kluster akhir pada pengujian ini dapat dilihat pada Tabel 7 berikut.

Tabel 7. Hasil kluster dengan nilai k=3

No.	Desa	PC1	PC2	Cluster
1	Suku Rawe	-3,777884863	0,047341326	1
2	Tanjung Sari	-2,048745529	0,23347204	1
3	Bintang Bener Mutiara	-5,089276896	-1,287932167	1
4	Oneng	-2,336452159	-0,353207761	1
5	Musara Ate	-4,022937196	0,223848408	1
6	Selamat Rejo	-3,465799934	1,081132877	1
7	Sedie Jadi	-4,172354545	0,229307284	1
---	---	---	---	---
---	---	---	---	---
30	Teleden	2,298915233	-0,174392418	3
31	Sumber Rejo	2,091001184	0,300667285	3
32	Bakongan	2,027132216	-0,080883232	3
33	Timur Jaya	2,287438401	-0,304643684	3
34	Reje Guru	2,776603123	-0,083823879	3

Hasil kluster dari tabel 7 juga dapat dilihat dengan bentuk visualisasi yang diplot menggunakan fungsi *ggplot()*, yang merupakan suatu fungsi dari *library(ggplot2)*. Visuasilasi kluster dengan nila k=3 dapat dilihat pada Gambar 3 berikut.



Gambar 3. Visualisasi kluster dengan nilai k=3

Pada visualisasi dapat dilihat bahwa 3 kluster ini dibedakan antara tiga warna yaitu merah, hijau dan biru, warna merah menunjukkan desa tersebut berada pada kluster 1 yaitu sebanyak 7 desa, warna hijau menunjukkan desa berada pada kluster 2 yaitu sebanyak 13 dan warna biru yang menunjukkan desa berada pada kluster 3 yaitu sebanyak 14.

4.3.3 Pengujian Kluster k=4

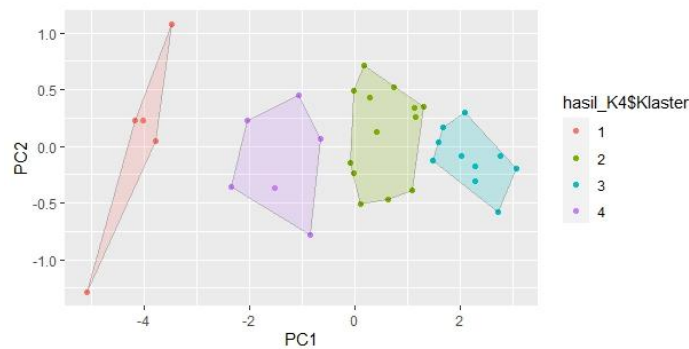
Pengujian pertama dilakukan dengan jumlah kluster sebanyak 4 kluster yang diuji terhadap data penjualan kopi sebanyak 34 desa. Adapun hasil kluster akhir pada pengujian ini dapat dilihat pada Tabel 8 berikut.

Tabel 8. Hasil kluster dengan nilai k=4

No	Desa	PC1	PC2	Cluster
1	Suku Rawe	-3,7779	0,04734	1

2	Bintang Bener Mutiara	-5,0893	-1,2879	1
3	Musara Ate	-4,0229	0,22385	1
4	Selamat Rejo	-3,4658	1,08113	1
5	Sedie Jadi	-4,1724	0,22931	1
6	Temas Mumanang	-0,0216	-0,234	2
7	Penosan Jaya	0,11351	-0,5051	2
---	---	---	---	---
---	---	---	---	---
27	Timur Jaya	2,28744	-0,3046	3
28	Reje Guru	2,7766	-0,0838	3
29	Tanjung Sari	-2,0487	0,23347	4
32	Batin Jaya	-0,8438	-0,7776	4
33	Waq Pondok Sayur	-1,0545	0,44711	4
34	Petukel	-0,6456	0,0676	4

Hasil kluster dari tabel 8 juga dapat dilihat dengan bentuk visualisasi yang diplot menggunakan fungsi *ggplot()*, yang merupakan suatu fungsi dari *library(ggplot2)*. Visualisasi kluster dengan nilai $k=4$ dapat dilihat pada Gambar 4 berikut.



Gambar 4. Visualisasi kluster dengan nilai $k=4$

Pada visualisasi dapat dilihat bahwa 4 kluster ini dibedakan antara empat warna yaitu merah, ungu, hijau dan biru, warna merah menunjukkan desa tersebut berada pada kluster 1 yaitu sebanyak 5 desa, warna hijau menunjukkan desa berada pada kluster 2 yaitu sebanyak 13 desa, warna biru yang menunjukkan desa berada pada kluster 3 yaitu sebanyak 10 desa dan warna ungu yang menunjukkan desa berada pada kluster 4 yaitu sebanyak 6 desa.

4.4 Hasil Validasi Kluster

Proses yang terakhir dari penelitian ini yaitu 34 data desa hasil klustering menggunakan metode *kharmonic means* kemudian divalidasi kluster yang sudah terbentuk dari mulai nilai $k=2$, $k=3$ dan $k=4$ yang akan dilakukan menggunakan validasi *silhouette index* dan *c-index*. Untuk menghitung nilai *silhouette index* juga dapat dilakukan menggunakan fungsi *silhouette()* pada *Rstudio* yang merupakan fungsi dari *library(cluster)* Dari Hasil proses perhitungan metode validasi kluster menggunakan metode *silhouette index* untuk semua jumlah kluster ini dapat dilihat pada Gambar 5 berikut.

	Kluster	Value
1	2	0.5982177
2	3	0.4969937
3	4	0.4630244

Gambar 5. Hasil validasi kluster *silhouette index*

Dari gambar diatas dapat kita lihat bahwa pada jumlah kluster 2 atau $k=2$ memiliki nilai sebesar 0.59, pada jumlah kluster 3 atau $k=3$ dengan nilai sebesar 0.49 dan yang terakhir pada jumlah

kluster 4 atau $k=4$ memiliki nilai sebesar 0.46. Selanjutnya yaitu melakukan validasi kluster menggunakan metode *C-Index*, perhitungan nilai pada metode tersebut dilakukan dengan menggunakan fungsi *c_index ()* pada *Rstudio*. Hasil dari nilai pada validasi kluster menggunakan metode *c-index* ini dapat dilihat pada Gambar 6 berikut.

	Kluster	Value
1	2	0.06635945
2	3	0.05207019
3	4	0.03436281

Gambar 6. Hasil validasi kluster *c-index*

Dari Gambar 6 di atas dapat kita lihat bahwa pada jumlah kluster 2 atau $k=2$ memiliki nilai *c-index* sebesar 0.06, pada jumlah kluster 3 atau $k=3$ dengan nilai *c-index* sebesar 0.05 dan yang terakhir pada jumlah kluster 4 atau $k=4$ memiliki nilai *c-index* sebesar 0.03.

4.5 Pembahasan Hasil Penelitian

Pada pembahasan hasil penelitian ini peneliti akan membahas hasil dari penelitian yang telah diperoleh pada sub bab sebelumnya, pada pembahasan ini akan membandingkan hasil validasi kluster yang dihasilkan dari dua metode validasi yaitu metode *silhouette index* dan *c-index*. Dalam pengambilan jumlah kluster yang optimal yaitu dengan melihat nilai *silhouette index* paling besar dan nilai *c-index* paling kecil dari percobaan jumlah kluster. Berdasarkan hasil pengujian tersebut dapat dilihat hasil perbandingan validasi kluster dari hasil kluster menggunakan metode *K-Harmonic Means* terhadap data penjualan kopi dari 34 desa sebagai berikut.

Tabel 9. Perbandingan hasil validasi kluster

Validasi Index	Jumlah Kluster		
	2	3	4
<i>Silhouette Index</i>	0,59822	0,496994	0,46302
<i>C-Index</i>	0,06635	0,052071	0,03436
<i>Silhouette Index</i> (Maksimum)	0,59822		
<i>C-Index</i> (Minimum)	0,03436		

Berdasarkan Tabel 9 menunjukkan bahwa, dari hasil uji validasi menggunakan *Silhouette Index* diperoleh kluster yang optimal adalah 2 kluster dengan nilai maksimum *Silhouette Index* sebesar 0,59, hal ini dikarenakan nilai tertinggi *Silhouette Index* terdapat pada kluster $k=2$ atau nilainya mendekati 1 diantara nilai yang lainnya. Sedangkan uji validasi pada metode *C-index* kluster yang optimal adalah 4 kluster dengan nilai *c-index* sebesar 0.03, yang artinya nilai minimum dari hasil *c-index* berada pada kluster 4 atau $k=4$.

4.5.1 Profiling Kluster $k=2$

Profiling kluster ini dilakukan pada 2 kluster terbaik yang sudah dilakukan validasi kluster menggunakan metode *Silhouette Index*. Untuk mendapatkan nilai profiling kluster yaitu menghitung rata-rata dari setiap variable berdasarkan kluster. Untuk hasil Profiling kluster $k=2$ dapat kita lihat pada Tabel 10 berikut

Tabel 10. Hasil Profiling kluster $k=2$

Kluster	Jumlah	Nama Desa	Kualitas Kopi
1	9	Suku Rawe, Tanjung Sari, Bener Lukup II, Bintang Bener Mutiara, Oneng, Musara Ate, Waq Pondok Sayur, Selamat Rejo dan Sedie Jadi	Sangat Baik
2	25	Temas Mumanang, Panji 1, Batin Jaya, Penosan Jaya, Wih Resap, Hakim Wih	Baik

Hilang, Bener Pepanyi, Ayu Ara, Delong Asli, Ujung Gele, Belang Tampu, Rembele Timur, Jadi Sepakat, Tingkem, Vitran Temas Mumanang, Petukel, Bukit Nosar, Teleden, Sumber Rejo, Tansaran Bidin, Kekuyang, Bakongan, Timur Jaya dan Reje Guru

Dari hasil profiling kluster $k=2$ diatas dapat kita lihat bahwa kluster 1 merupakan kelompok desa dengan kualitas kopi sangat baik, dengan jumlah desa sebanyak 9 desa. Sedangkan kluster 2 menunjukkan kelompok desa dengan kualitas kopi baik dengan jumlah desa sebanyak 25 desa.

4.5.2 Profiling Kluster $k=4$

Profiling kluster ini dilakukan pada 4 kluster terbaik yang sudah dilakukan validasi kluster menggunakan metode *C-Index*. Untuk mendapatkan nilai profiling kluster yaitu menghitung rata-rata dari setiap variable berdasarkan kluster. Untuk hasil Profiling kluster $k=4$ dapat kita lihat pada Tabel 11 berikut.

Tabel 11. Hasil Profiling kluster $k=4$

Kluster	Jumlah	Nama Desa	Kualitas Kopi
1	5	Suku Rawe, Bintang Bener Mutiara, Musara Ate, Selamat Rejo dan Sedie Jadi	Sangat Baik
2	13	Temas Mumanang, Penosan Jaya, Wih Resap, Bener Pepanyi, Ayu Ara, Delong Asli, Ujung Gele, Jadi Sepakat, Vitran Temas Mumanang, Bukit Nosar, Wonosari, Tansaran Bidin dan Kekuyang	Baik
3	10	Panji 1, Hakim Wih Hilang, Belang Tampu, Rembele Timur, Tingkem, Teleden, Sumber Rejo, Bakongan, Timur Jaya dan Reje Guru	Cukup
4	6	Tanjung Sari, Bener Lukup II, Oneng, Batin Jaya, Waq Pondok Sayur dan Petukel	Kurang

Dari hasil profiling kluster $k=4$ diatas dapat kita lihat bahwa kluster 1 merupakan kelompok desa dengan kualitas kopi sangat baik, dengan jumlah desa sebanyak 5 desa. Kluster 2 menunjukkan kelompok desa dengan kualitas kopi baik dengan jumlah desa sebanyak 13 desa. Kluster 3 menunjukkan kelompok desa dengan kualitas kopi cukup dengan jumlah desa sebanyak 10 desa. Sedangkan pada kluster 4 menunjukkan kelompok desa dengan kualitas kopi kurang dengan jumlah desa sebanyak 6 desa.

4.6 Pemilihan Metode Validasi Kluster Terbaik

Untuk menentukan metode validasi kluster yang terbaik dari metode *silhouette index* dan *c-index* dapat digunakan metode salah satunya yaitu menggunakan rata-rata simpangan baku. Selanjutnya untuk menentukan metode validasi yang terbaik dapat menggunakan rata-rata simpangan baku dalam kluster (S_w) dan rata-rata simpangan baku antar kluster (S_b). Metode yang terbaik dapat dilihat dari nilai rasio yang terkecil dari sw/sb . Adapun hasil perhitungan sw/sb dari validasi kluster *silhouette index* dan *c-index* dapat dilihat pada Tabel 12 berikut.

Tabel 12. Hasil validasi jumlah kluster

Metode Validasi	Kluster	S_w	S_b	S_w/S_b
<i>Silhouette Index</i>	2	10.571	90.423	0,1169
<i>C-Index</i>	4	6.124	83.454	0,0734

Dari hasil tersebut jumlah kluster 2 dengan validasi kluster *Silhouette Index* memiliki nilai rata-rata simpangan baku sebesar 0.116, sedangkan pada validasi kluster *C-Index* memiliki nilai rata-rata simpangan baku sebesar 0.073, kedua nilai rasio ini yang terkecil terdapat pada validasi kluster *C-*

Index, dan dapat dikatakan bahwa klastering kualitas kopi berdasarkan taraf penyusutan paling optimal yaitu menggunakan 4 klaster.

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan terhadap klastering kualitas biji kopi dari 34 desa berdasarkan taraf penyusutan mendapatkan Hasil klastering menggunakan metode *c-index* klaster yang terbaik yaitu sebanyak 4 klaster dengan nilai $CI=0,03436$ menghasilkan 5 desa dengan kategori kualitas kopi sangat baik, 13 desa dengan kategori kualitas kopi baik, 10 desa dengan kualitas kopi cukup dan 6 desa dengan kualitas kurang sedangkan hasil klastering menggunakan metode *Silhouette Index* klaster yang terbaik yaitu sebanyak 2 klaster dengan nilai $SI=0,59822$ terdapat 9 desa pada dengan kategori kualitas kopi sangat baik dan 25 desa dengan kategori kualitas kopi baik. Adapun hasil klaster dari kedua metode tersebut yang paling optimal yaitu menggunakan 4 klaster, yang telah divalidasi menggunakan metode *c-index*, karena pada metode ini memiliki nilai rasio Sw dan Sb terkecil yaitu 0,0734. Untuk penelitian selanjutnya diharapkan dapat dilakukan validasi klaster menggunakan metode validasi *index* lainnya dan dapat menggunakan data yang terbaru dan dapat menambah variabel lain yang berhubungan dengan data tersebut serta karena masih ada beberapa desa yang masuk ke dalam kategori kualitas kopi yang kurang baik, maka untuk pihak yang terkait dapat melakukan evaluasi terhadap desa-desa tersebut agar sama dengan desa-desa yang lain.

DAFTAR PUSTAKA

- [1] Y. Putriana, A. Amri, and H. T. Hidayat, "Penentuan Kualitas Biji Kopi Menggunakan Metode Fuzzy Logic," *J. Infomedia*, vol. 1, no. 1, 2016, doi: 10.30811/v1i1.280.
- [2] E. T. Kembaren and Muchsin, "Pengelolaan Pasca Panen Kopi Arabika Gayo Aceh," *J. Visioner dan Strateg.*, vol. 10, no. 1, pp. 29–36, 2021.
- [3] "BPS Bener Meriah," 2020. <https://benermeriahkab.bps.go.id/> (accessed Jan. 19, 2022).
- [4] S. 'Aina Salsabila, T. Widiari, and Sudarno, "Metode K-Harmonic Means Clustering dengan Validasi Silhouette Coefficient," *J. Gaussian*, vol. 11, no. 1, pp. 11–20, 2022.
- [5] Y. D. Gustientiedina, M. Hasmil Adiya, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [6] D. F. Pasariibu, I. S. Damanik, E. Irawan, Suhada, and H. S. Tambunan, "Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 11–20, 2021, doi: 10.37148/bios.v2i1.17.
- [7] M. A. W. K. MURTI, "Penerapan Metode K-Means Clustering Untuk Mengelompokan Potensi Produksi Buah – Buahan Di Provinsi Daerah Istimewa Yogyakarta," *Skripsi*, 2017.
- [8] D. Abdullah, S. Susilo, A. S. Ahmar, R. Rusli, and R. Hidayat, "The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data," *Qual. Quant.*, vol. 56, no. 3, pp. 1283–1291, 2022, doi: 10.1007/s11135-021-01176-w.
- [9] I. gede Gunadi, D. P. S. Putri, I. . Mistanada, G. . Prayoga, and N. M. Y. . Rahayu, "Analisis Cluster Pada Pengelompokan Siswa Diktuk Bintara Polri TA. 2018/2019, SPN Singaraja - Polda Bali Menggunakan K-Means dan K-Harmonic Means," *J. Ilm. SINUS*, vol. 17, no. 2, p. 13, 2019, doi: 10.30646/sinus.v17i2.421.
- [10] Y. Syahra, R. I. Ginting, and M. Yetri, "Implementasi Data Mining Untuk Penyusunan Tata Letak Data Obat-Obatan Dengan Menggunakan Algoritma K-Harmonic Means Pada Apotek Inti Fada Sidamanik," *J. Technopreneursh. Inf. Syst.*, vol. 3, no. 1, pp. 19–25, 2020, doi: 10.36085/jtis.v3i1.718.
- [11] M. A. Nahdliyah, T. Widiari, and A. Prahutama, "Metode K-Medois Clustering Dengan Validasi Silhouette Index Dan C-Index (Studi Kasus Jumlah Kriminalitas Kabupaten/Kota di Jawa Tengah Tahun 2018)," *J. Gaussian*, vol. 8, no. 2, pp. 161–170, 2019, doi: 10.14710/j.gauss.v8i2.26640.
- [12] M. Charrad, N. Ghazzali, V. Boiteau, and A. Niknafs, "Nbclust: An R package for determining the relevant number of clusters in a data set," *J. Stat. Softw.*, vol. 61, no. 6, pp. 1–36, 2014, doi: 10.18637/jss.v061.i06.
- [13] H. H. Qastari, O. Soesanto, and Y. Sukmawaty, "K-Means Clustering dan Principal Component Analysis (PCA) Dalam Radial Basis Function Neural Network (RBFNN) Untuk Klasifikasi Data Multivariat," *J. Math. Theory Appl.*, vol. 4, no. 1, pp. 1–7, 2022, doi: 10.31605/jomta.v4i1.1757.
- [14] D. I. Yunistya, R. Goejantoro, and F. D. T. Amijaya, "The Application Of K – Harmonic Means Method In District / City Grouping (Case Study : Poverty in Kalimantan Island in 2020)," *J. Mat. Stat. dan komputasi*, vol. 19, no. 1, pp. 51–64, 2022, doi: 10.20956/j.v19i1.21116.
- [15] B. Poerwanto, "Evaluating the K-Means Analysis in Clustering Area Based on Estates Productivity in

- Tana Luwu Using Silhouette Index,” *J. Phys. Conf. Ser.*, vol. 1752, no. 1, 2021, doi: 10.1088/1742-6596/1752/1/012014.
- [16] H. Mahi, N. Farhi, K. Labed, and D. Benhamed, “The Silhouette Index and the K-Harmonic Means algorithm for Multispectral Satellite Images Clustering,” *Proc. 2018 Int. Conf. Appl. Smart Syst. ICASS 2018*, no. November, pp. 1–6, 2019, doi: 10.1109/ICASS.2018.8652068.
- [17] A. Assegaf, M. A. Mukid, and A. Hoyyi, “Analisis Kesehatan Bank Menggunakan Local Mean K-Nearest Neighbor dan Multi Local Means K-Harmonic Nearest Neighbor,” *J. Gaussian*, vol. 8, no. 3, pp. 343–355, 2019, doi: 10.14710/j.gauss.v8i3.26679.
- [18] T. P. Fiqar, S. B. Musa, F. M. Humaira, I. M. Widiartha, D. Herumurti, and A. Z. Arifin, “Pengembangan Metode Klasterisasi Data Berbasis Hybrid Improved Artificial Bee Colony (IABC) dan K – Harmonic Means,” *SPECTA J. Technol.*, vol. 2, no. 3, p. 10, 2018, doi: 10.35718/specta.v2i3.3.