

**PENERAPAN *CLUSTERING TIME SERIES* UNTUK MENGELOMPOKKAN
PROVINSI DI INDONESIA BERDASARKAN JUMLAH PRODUKSI KOPI
TAHUN 2008-2021**

**Nadya Pratiwi Doholio¹, Rifindu Putra Oktafian², Dian Febrianti Jafar³,
Prycillia M. C. Nelwan⁴, Moh. Adyan Balada⁵, Asrul S. Ulopo⁶, Stevani
Hanapi⁷, Sarah Aulia Djafar⁸, Rahmat Septian Daud⁹, La Ode Nashar¹⁰**

Universitas Negeri Gorontalo

¹nadyapратиwi39@gmail.com, ²rifinduputra751@gmail.com,
³dianjafar04@gmail.com, ⁴nprycillia@gmail.com, ⁵adyanbalada@gmail.com,
⁶asrululopo@gmail.com, ⁷stevanihanapi77@gmail.com,
⁸alyadjafar80@gmail.com, ⁹septiandaud567@gmail.com,
¹⁰laode.nashar@ung.ac.id

Received 21 October 2022; revised 11 November 2022; accepted 9 December 2022.

ABSTRAK

Kopi merupakan salah satu komoditas perkebunan di Indonesia yang peranannya begitu penting. Indonesia terdapat pada posisi keempat sebagai negara penghasil kopi terbesar. Hal ini bisa saja dikarenakan wilayah Indonesia yang memungkinkan untuk memproduksi tanaman kopi lebih banyak. Karakteristik wilayah di Indonesia berbeda-beda, sehingga penghasilan tanaman kopi juga berbeda-beda setiap tahunnya. Tujuan diadakannya penelitian ini adalah untuk mengelompokkan provinsi-provinsi di Indonesia dengan melakukan penerapan *Clustering Time Series* pada data nilai produksi kopi setiap provinsi di Indonesia pada periode tahun 2008-2021. Jarak terbaik yang dapat digunakan untuk *Clustering Time Series Analysis* pada data ini yaitu jarak *Dynamic Time Warping* (DTW) dengan metode *Average Linkage* dan menghasilkan 3 *cluster*, yaitu *cluster* 1 terdiri dari 2 provinsi, *cluster* 2 terdiri dari 4 provinsi, dan *cluster* 3 terdiri dari 28 provinsi. Secara rata-rata, 3 *cluster* yang terbentuk menghasilkan nilai koefisien *silhouette* 0.81 atau dikategorikan *Strong Classification*.

Kata kunci: deret waktu, cluster, kopi.

ABSTRACT

Coffee is one of the most important plantation commodities in Indonesia. Indonesia is in the fourth position as the largest coffee producing country. This could be due to the area of Indonesia that allows to produce more coffee plants. The characteristics of regions in Indonesia are

different, so the income of coffee plants also varies every year. The purpose of this study is to group provinces in Indonesia by applying Time Series Clustering to coffee production data for each province in Indonesia for the period 2008-2021. The best distance that can be used for Clustering Time Series Analysis on this data is the Dynamic Time Warping (DTW) distance with the Average Linkage method and produces 3 clusters, namely cluster 1 consisting of 2 provinces, cluster 2 consisting of 4 provinces, and cluster 3 consisting of 28 provinces. On average, the 3 clusters formed produce a silhouette coefficient value of 0.81 or categorized as Strong Classification.

Keywords: time series, cluster, coffee.

PENDAHULUAN

Sektor perkebunan menjadi salah satu sektor penting dalam perekonomian Indonesia. Sektor perkebunan merupakan salah satu sumber pangan dunia, membuat perkebunan di Indonesia menjadi sangat penting. Ditambah lagi perkebunan menjadi salah satu komoditas ekspor Indonesia yang menjadi penopang perekonomian negara. Sektor perkebunan juga tidak hanya menjadi sumber perekonomian, namun menjadi sumber tenaga kerja Indonesia yang tidak sedikit masyarakatnya berprofesi petani. Sektor perkebunan merupakan bagian dari sektor pertanian yang meningkatkan deviasi negara serta menyerap tenaga kerja. Pemerintah mengutamakan pada subsektor perkebunan, karena subsektor ini memiliki daya tarik yang tinggi untuk diekspor ke negara maju (Soediono, 1989).

Kopi merupakan salah satu komoditas perkebunan Indonesia yang memiliki peran penting terutama sebagai sumber devisa dan juga penyedia lapangan kerja sekaligus sumber pendapatan bagi pelaku ekonomi yang berkaitan dengan kopi. Perdagangan kopi di Indonesia menjadi cukup besar.

Berdasarkan data dari *International Coffee Organization* (2017), Indonesia terdapat pada posisi keempat sebagai negara penghasil kopi terbesar dari seluruh negara produsen kopi di dunia. Hal ini menunjukkan negara kita Indonesia mempunyai potensi besar dalam komoditas kopi, selain sebagai produsen juga sebagai negara eksportir kopi terbesar posisi keempat di dunia.

Proses produksi kopi merupakan kegiatan manusia dengan melibatkan sumber daya lainnya (metode, mesin, modal) secara produktif dalam mengolah

kopi menjadi produk kopi yang dapat dijual ke konsumen (tidak perlu). Produksi kopi di Indonesia dapat dilakukan hampir di seluruh wilayah Indonesia, kecuali di wilayah yang memiliki suhu lingkungan tinggi.

Karakteristik wilayah yang berbeda-beda dapat mempengaruhi nilai produksi kopi yang dihasilkan. Karakteristik masing-masing wilayah dapat diidentifikasi dengan menggunakan analisis gerombol (*Cluster Analysis*). Analisis ini perlu dilakukan dengan tujuan agar program kebijakan yang diadakan pemerintah untuk Indonesia terkait sektor perkebunan dapat lebih berfokus dan terarah sesuai dengan karakteristik dari masing-masing *cluster* yang terbentuk.

Analisis *Cluster* adalah salah satu metode dalam statistika yang memungkinkan peneliti untuk mengelompokkan atau mengklasifikasikan objek ke dalam kelompok yang relatif homogen. Objek atau kasus dalam setiap kelompok cenderung memiliki karakteristik yang mirip satu sama lain dan memiliki perbedaan dengan objek dari kelompok lainnya. Dengan kata lain, keragaman dalam suatu kelompok atau gerombol atau *cluster* menjadi minimum sedangkan keragaman antar kelompok atau gerombol atau *cluster* menjadi maksimum (Mattjik dan Sumertajaya, 2011).

Penggunaan analisis *cluster* semakin berkembang dengan adanya penggunaan pada data runtun waktu (*time series*). Kumpulan data *time series* dapat dikelompokkan berdasarkan dengan karakteristik pada masing-masing runtun waktu atau deret waktu. Penerapan analisis *cluster* runtun waktu (*clustering time series analysis*) dilakukan untuk mengelompokkan objek berdasarkan pola deret waktunya. Tetapi pemilihan jarak dan metode *clustering* yang digunakan harus bisa mewadahi struktur data *time series* yang sifatnya dinamis (Liao, 2005).

Berdasarkan latar belakang di atas, maka penelitian ini dilakukan dengan tujuan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan nilai produksi kopi dengan melakukan analisis *clustering* runtun waktu (*clustering time series analysis*). Penelitian ini akan dilakukan berdasarkan data produksi kopi pada masing-masing provinsi di Indonesia pada periode tahun 2008-2021.

METODE PENELITIAN

Dalam penelitian ini menggunakan metode penelitian kuantitatif. Menurut V. Wiratna Sujarweni (2014), penelitian kuantitatif merupakan penelitian yang dapat menghasilkan penemuan yang diperoleh dengan menggunakan prosedur-prosedur statistik atau cara lain dari pengukuran (kuantifikasi).

Sumber data yang digunakan untuk penelitian ini yaitu data sekunder yang bersumber dari data Badan Pusat Statistik (www.bps.go.id). Data yang dimaksud yakni data Nilai Produksi Kopi di Provinsi Indonesia dari tahun 2008 sampai 2021. Menurut Sugiyono (2018), data sekunder merupakan sumber data yang tidak langsung dikumpulkan oleh peneliti, misalnya melalui orang lain atau melalui dokumen. Pada penelitian ini sumber data yaitu melalui dokumen.

Metode yang digunakan untuk analisis data yakni Analisis *Cluster Time Series* (*Clustering Time Series Analysis*). Analisis *Cluster* merupakan sebuah teknik dengan tujuan menemukan suatu kelompok dalam kumpulan data, agar data dalam satu kelompok sangat mirip dan juga sangat berbeda secara signifikan dengan kelompok yang lain (Putu Widya Adnyani & Robinson Sihombing, 2021). Analisis *Cluster Time Series* membutuhkan algoritma dan prosedur pengelompokan yang berbeda, hal ini karena data deret waktu merupakan sekelompok nilai pengamatan yang diperoleh pada waktu yang berbeda dengan selang waktu yang sama serta barisan data diasumsikan saling berhubungan satu sama lain. Selama perkembangannya, banyak teknik yang dapat dikembangkan pada data deret waktu, diantaranya yaitu penggunaan jarak-jarak pengelompokan yang sesuai dengan karakteristik deret waktu (Liao, 2005).

Langkah-langkah untuk menganalisis data yang dilakukan pada penelitian ini yaitu:

- a. Eksplorasi data.
- b. Melakukan *clustering* metode hierarki analisis *cluster time series* dengan menggunakan ukuran jarak *Dynamic Time Warping* (DTW) dan *Euclidean*. Jarak DTW merupakan sebuah algoritma yang bertujuan menemukan jalur *warp* yang optimal antara dua data *time series* sehingga menghasilkan kumpulan nilai jalur *warp* dan jarak antara dua deret waktu (*time series*). Jarak DTW merupakan ukuran ketaksamaan yang melibatkan prediksi tanpa model

tertentu serta merupakan jarak terkecil diantara dua pasang titik mengingat kemungkinan migrasi titik (Putu Widya Adnyani & Robinson Sihombing, 2021).

Lintasan *warping* pada penggunaan jalur DTW ini menggunakan jarak *Euclidean*. Jalur *warping* yang memiliki jarak minimum antara dua deret waktu memiliki rumus yaitu:

$$d_{DTW} = \min \frac{\sum_{k=1}^K w_k}{K}$$

dengan:

d_{DTW} : jarak *Dynamic Time Warping*

w_k : lintasan *warping* ke-k

Perhitungan jarak DTW menggunakan 3 metode keterpautan yakni "*single linkage*", "*complete linkage*" dan "*average linkage*".

Jarak *Euclidean* merupakan salah satu jenis pengukuran jarak dalam analisis *cluster* yang akan mengukur jarak dari objek data ke pusat *cluster*. Jarak *Euclidean* ini merupakan jarak geometris antara dua objek data. Jika datanya semakin kecil, maka akan semakin mirip objek datanya (Dani et al., 2019). Perhitungan jarak *Euclidean* dalam penelitian ini juga akan menggunakan 3 metode keterpautan yakni "*single linkage*", "*complete linkage*" dan "*average linkage*".

- c. Melakukan perhitungan korelasi *Cophenetic*. Koefisien korelasi *Cophenetic* adalah koefisien korelasi antara elemen-elemen asli matriks ketakmiripan dan elemen-elemen yang dihasilkan oleh matriks *Cophenetic* (Silva & Dias, 2013).

Rumus yang digunakan untuk menghitung koefisien ini yaitu:

$$r_{Coph} = \frac{\sum_{i < k} (d_{ik} - \bar{d})(d_{cik} - \bar{d}_c)}{\sqrt{[\sum_{i < k} (d_{ik} - \bar{d})^2][\sum_{i < k} (d_{cik} - \bar{d}_c)^2]}}$$

dengan:

r_{Coph} : koefisien korelasi *Cophenetic*

d_{ik} : jarak *Euclidean* objek ke-i dan ke-k

$\bar{d}$: rata-rata d_{ik}

d_{cik} : jarak *Cophenetic* objek ke-i dan ke-k

$\bar{d}_c$: rata-rata d_{cik}

Nilai dari koefisien korelasi *Cophenetic* ini hanya berada diantara -1 dan 1. Jika nilai semakin mendekati 1, maka berarti bahwa solusi yang dihasilkan dari proses *clustering* semakin baik. Ukuran jarak DTW ataupun jarak *Euclidean* dengan metode keterhubungan yang dapat menghasilkan nilai *cophenetic* terbesar merupakan solusi terbaik pada metode *clustering* hierarki.

- d. Menghitung koefisien *silhouette* untuk melihat ketepatan *clustering*.

Rumus untuk menghitung koefisien ini yaitu:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

dimana:

$a(i)$: rata-rata jarak antar anggota dalam *cluster*

$b(i)$: jarak terkecil antara anggota *cluster* dengan anggota *cluster* terdekatnya (*nearest neighbor*).

Kriteria yang digunakan untuk melihat ketepatan dan kualitas hasil *clustering* berdasarkan nilai koefisien *silhouette* ini adalah sebagai berikut:

Tabel 1. Kategori *Clustering* berdasarkan Koefisien *Silhouette*

Nilai Koefisien <i>Silhouette</i> (SC)	Kategori <i>Clustering</i>
0.71 – 1.00	<i>Strong Classification</i>
0.51 – 0.70	<i>Good Classification</i>
0.26 – 0.50	<i>Weak Classification</i>
0 – 0.25	<i>Bad Classification</i>

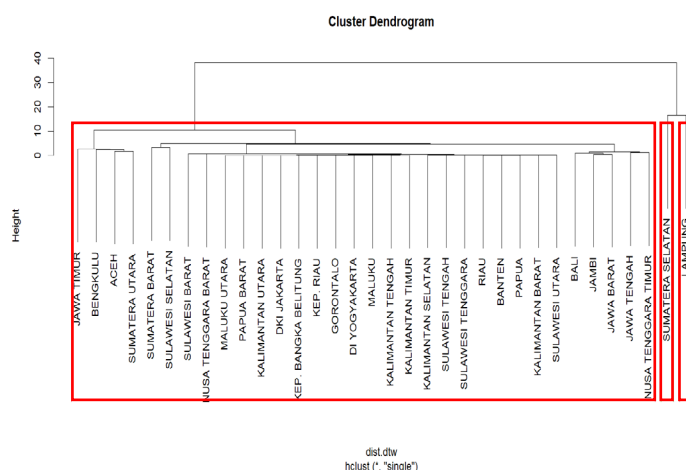
Sumber: Kauffman dan Rousseuw (1990)

HASIL PENELITIAN DAN PEMBAHASAN

Langkah pertama adalah menggambarkan dendogram hasil analisis *clustering time series* dengan metode hierarki menggunakan jarak DTW dengan metode keterpautannya (*single linkage*, *complete linkage*, *average linkage*). Hasilnya adalah sebagai berikut.

- a. *Single Linkage*

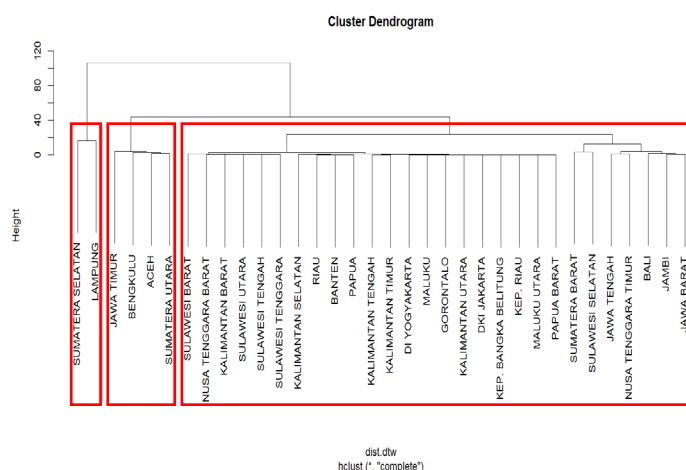
Penerapan Clustering Time Series untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Jumlah Produksi Kopi Tahun 2008-2021



Gambar 1. Single Linkage Jarak DTW

Berdasarkan dendrogram di atas dapat dilihat bahwa terbentuk tiga kelompok atau tiga *cluster* yang terdiri dari *cluster* pertama sebanyak 32 provinsi yaitu Jawa Timur, Bengkulu, Aceh, Sumatera Utara, Sumatera Barat, Sulawesi Selatan, Sulawesi Barat, NTB, Maluku Utara, Papua Barat, Kalimantan Utara, DKI Jakarta, Kep. Bangka Belitung, Kep. Riau, Gorontalo, DI Yogyakarta, Maluku, Kalimantan Tengah, Kalimantan Timur, Kalimantan Selatan, Sulawesi Tengah, Sulawesi Tenggara, Riau, Banten, Papua, Kalimantan Barat, Sulawesi Utara, Bali, Jambi, Jawa Barat, Jawa Tengah, dan NTT. Kemudian *cluster* kedua sebanyak 1 provinsi yaitu Sumatera Selatan. Selanjutnya yang terakhir *cluster* ketiga sebanyak 1 provinsi yaitu Lampung.

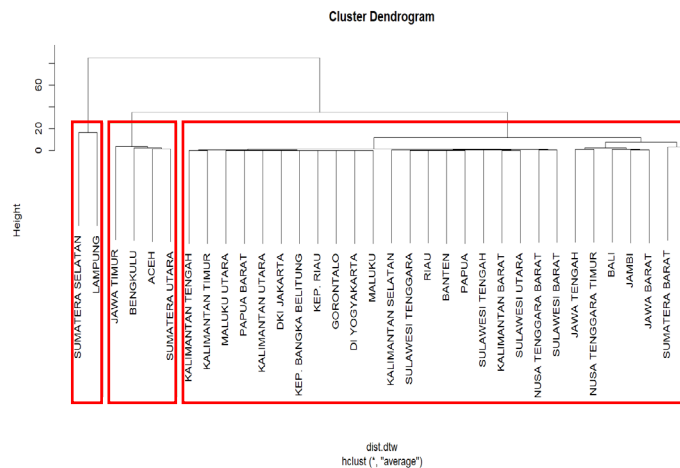
b. Complete Linkage



Gambar 2. Complete Linkage Jarak DTW

Berdasarkan dendrogram di atas dapat dilihat bahwa terbentuk tiga kelompok atau tiga *cluster* yang terdiri dari *cluster* pertama sebanyak 2 provinsi yaitu Sumatera Selatan dan Lampung. Kemudian *cluster* kedua sebanyak 4 provinsi yaitu Jawa Timur, Bengkulu, Aceh, dan Sumatera Utara. Selanjutnya yang terakhir *cluster* ketiga sebanyak 28 provinsi yaitu Sulawesi Barat, NTB, Kalimantan Barat, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Kalimantan Selatan, Riau, Banten, Papua, Kalimantan Tengah, Kalimantan Timur, DI Yogyakarta, Maluku, Gorontalo, Kalimantan Utara, DKI Jakarta, Kep. Bangka Belitung, Kep. Riau, Maluku Utara, Papua Barat, Sumatera Barat, Sulawesi Selatan, Jawa Tengah, NTT, Bali, Jambi, dan Jawa Barat.

c. *Average Linkage*



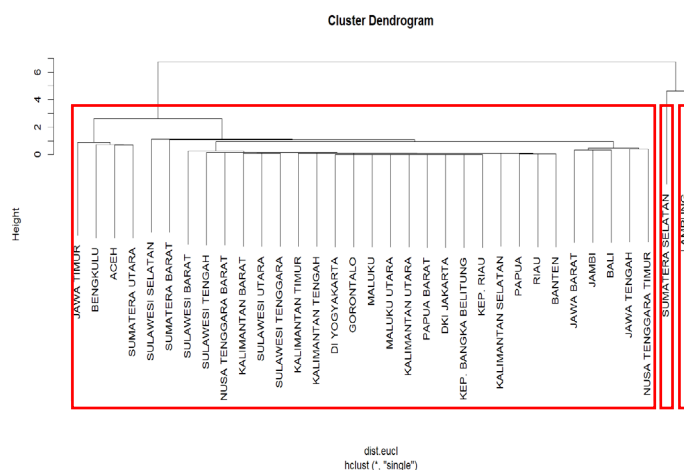
Gambar 3. *Average Linkage* Jarak DTW

Berdasarkan dendrogram di atas dapat dilihat bahwa terbentuk tiga kelompok atau tiga *cluster* yang terdiri dari *cluster* pertama sebanyak 2 provinsi yaitu Sumatera Selatan dan Lampung. Kemudian *cluster* kedua sebanyak 4 provinsi yaitu Jawa Timur, Bengkulu, Aceh, dan Sumatera Utara. Selanjutnya yang terakhir *cluster* ketiga sebanyak 28 provinsi yaitu Kalimantan Tengah, Kalimantan Timur, Maluku Utara, Papua Barat, Kalimantan Utara, DKI Jakarta, Kep. Bangka Belitung, Kep. Riau, Gorontalo, DI Yogyakarta, Maluku, Kalimantan Selatan, Sulawesi Tenggara, Riau, Banten, Papua, Sulawesi Tengah, Kalimantan Barat, Sulawesi Utara, NTB, Sulawesi Barat, Jawa Tengah, NTT, Bali, Jambi, Jawa Barat, Sumatera Barat, dan Sulawesi Selatan.

Penerapan Clustering Time Series untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Jumlah Produksi Kopi Tahun 2008-2021

Langkah selanjutnya adalah menggambarkan dendrogram hasil analisis *clustering time series* dengan metode hierarki menggunakan jarak *Euclidean* dan penerapan metode keterpautannya (*single linkage*, *complete linkage*, *average linkage*). Hasilnya adalah sebagai berikut.

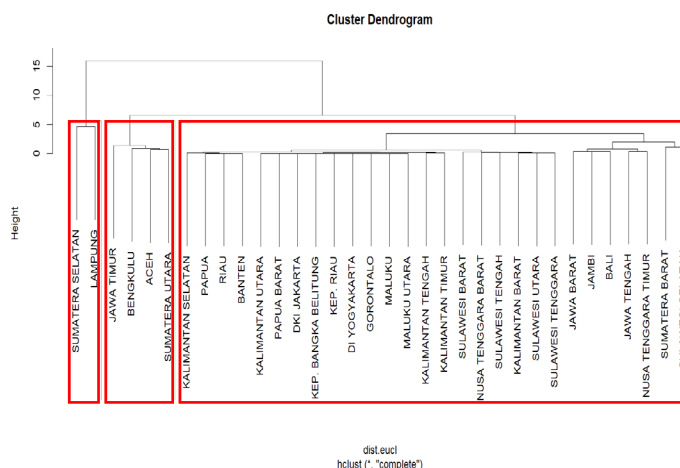
a. *Single Linkage*



Gambar 4. *Single Linkage* Jarak *Euclidean*

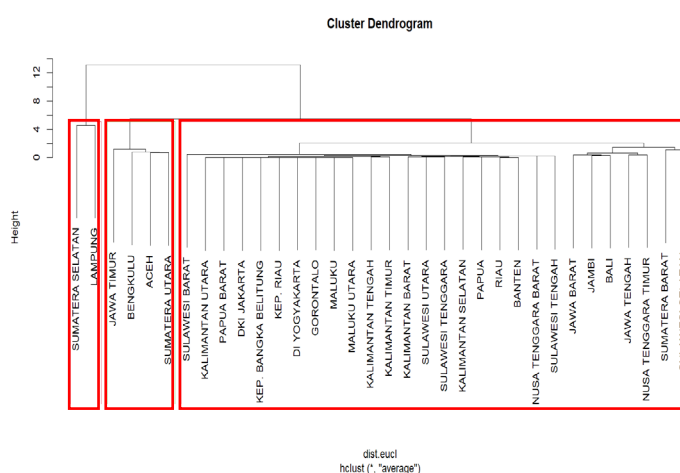
Berdasarkan dendrogram di atas dapat dilihat bahwa terbentuk tiga kelompok atau tiga *cluster* yang terdiri dari *cluster* pertama sebanyak 32 provinsi yaitu Jawa Timur, Bengkulu, Aceh, Sumatera Utara, Sulawesi Selatan, Sumatera Barat, Sulawesi Barat, Sulawesi Tengah, NTB, Kalimantan Barat, Sulawesi Utara, Sulawesi Tenggara, Kalimantan Timur, Kalimantan Tengah, DI Yogyakarta, Gorontalo, Maluku, Maluku Utara, Kalimantan Utara, Papua Barat, DKI Jakarta, Kep. Bangka Belitung, Kep. Riau, Kalimantan Selatan, Papua, Riau, Banten, Jawa Barat, Jambi, Bali, Jawa Tengah, dan NTT. Kemudian *cluster* kedua sebanyak 1 provinsi yaitu Sumatera Selatan. Selanjutnya yang terakhir *cluster* ketiga sebanyak 1 provinsi yaitu Lampung.

b. *Complete Linkage*

Gambar 5. *Complete Linkage Jarak Euclidean*

Berdasarkan dendrogram di atas dapat dilihat bahwa terbentuk tiga kelompok atau tiga *cluster* yang terdiri dari *cluster* pertama sebanyak 2 provinsi yaitu Sumatera Selatan dan Lampung. Kemudian *cluster* kedua sebanyak 4 provinsi yaitu Jawa Timur, Bengkulu, Aceh, dan Sumatera Utara. Selanjutnya yang terakhir *cluster* ketiga sebanyak 28 provinsi yaitu Kalimantan Selatan, Papua, Riau, Banten, Kalimantan Utara, Papua Barat, DKI Jakarta, Kep. Bangka Belitung, Kep. Riau, DI Yogyakarta, Gorontalo, Maluku, Maluku Utara, Kalimantan Tengah, Kalimantan Timur, Sulawesi Barat, NTB, Sulawesi Tengah, Kalimantan Barat, Sulawesi Utara, Sulawesi Tenggara, Jawa Barat, Jambi, Bali, Jawa Tengah, NTT, Sumatera Barat, dan Sulawesi Selatan.

c. *Average Linkage*

Gambar 6. *Average Linkage Jarak Euclidean*

Berdasarkan dendrogram di atas dapat dilihat bahwa terbentuk tiga kelompok atau tiga *cluster* yang terdiri dari *cluster* pertama sebanyak 2 provinsi yaitu

Sumatera Selatan dan Lampung. Kemudian *cluster* kedua sebanyak 4 provinsi yaitu Jawa Timur, Bengkulu, Aceh, dan Sumatera Utara. Selanjutnya yang terakhir *cluster* ketiga sebanyak 28 provinsi yaitu Sulawesi Barat, Kalimantan Utara, Papua Barat, DKI Jakarta, Kep. Bangka Belitung, Kep. Riau, DI Yogyakarta, Gorontalo, Maluku, Maluku Utara, Kalimantan Tengah, Kalimantan Timur, Kalimantan Barat, Sulawesi Utara, Sulawesi Tenggara, Kalimantan Selatan, Papua, Riau, Banten, NTB, Sulawesi Tengah, Jawa Barat, Jambi, Bali, Jawa Tengah, NTT, Sumatera Barat, dan Sulawesi Selatan.

Langkah selanjutnya yang perlu dilakukan yaitu menghitung nilai korelasi *cophenetic*, hasil yang diperoleh seperti yang ditunjukkan pada Tabel 2 berikut.

Tabel 2. Nilai Korelasi <i>Cophenetic</i>			
Jarak	Metode Keterpautan		
	<i>Single</i>	<i>Complete</i>	<i>Average</i>
<i>Euclidean</i>	0.9472	0.9557	0.9616
DTW	0.9697	0.9679	0.9711

Sumber: Hasil olah data

Dari tabel diatas, dapat dilihat bahwa setiap jarak dengan metode keterpautan yang berbeda mempunyai nilai koefisien korelasi *cophenetic* berbeda. Nilai koefisien korelasi *cophenetic* tertinggi adalah dengan menggunakan jarak DTW pada ketiga metode keterpautan tersebut. Maka dapat disimpulkan bahwa proses pengelompokkan provinsi di Indonesia berdasarkan nilai produksi kopi tahun 2008-2021, jarak pengukuran terbaik yang dapat digunakan yaitu *Dynamic Time Warping* (DTW).

Adapun metode keterpautan terbaik diantara ketiga metode keterpautan yang digunakan adalah *average linkage*, karena memiliki nilai koefisien korelasi *cophenetic* terbesar yakni 0.9711.

Distribusi provinsi-provinsi pada 3 *cluster* sesuai metode keterpautan *average linkage* dengan jarak DTW yakni:

- (1) *Cluster* pertama terdiri dari 2 provinsi yaitu Sumatera Selatan dan Lampung.
- (2) *Cluster* kedua terdiri dari 4 provinsi yaitu Jawa Timur, Bengkulu, Aceh, dan Sumatera Utara.
- (3) *Cluster* ketiga terdiri dari 28 provinsi yaitu Kalimantan Tengah, Kalimantan Timur, Maluku Utara, Papua Barat, Kalimantan Utara, DKI Jakarta, Kep.

Bangka Belitung, Kep. Riau, Gorontalo, DI Yogyakarta, Maluku, Kalimantan Selatan, Sulawesi Tenggara, Riau, Banten, Papua, Sulawesi Tengah, Kalimantan Barat, Sulawesi Utara, NTB, Sulawesi Barat, Jawa Tengah, NTT, Bali, Jambi, Jawa Barat, Sumatera Barat, dan Sulawesi Selatan.

Setelah memperoleh jarak terbaik yaitu jarak DTW serta metode terbaik yaitu *average linkage* dan menghasilkan 3 *cluster*, selanjutnya akan dilakukan analisis untuk menguji apakah jumlah *cluster* tersebut dapat mewakili atau representatif. Uji yang digunakan untuk menguji kualitas *cluster* ini yakni dengan koefisien *silhouette*.

Pada hasil olah data, nilai koefisien *silhouette* pada jumlah *cluster* sama dengan 3 yaitu sebesar 0.81. Nilai tersebut jika dilihat pada tabel 1 kategori *clustering* berdasarkan koefisien *silhouette*, menunjukkan bahwa pengelompokkan menggunakan metode *average linkage* dengan jarak DTW dan membentuk 3 *cluster* dengan kategori "tinggi", "sedang", "rendah" termasuk ke dalam kategori *clustering strong classification*.

SIMPULAN

Analisis *cluster time series* dalam penelitian ini memiliki tujuan untuk mengelompokkan provinsi di Indonesia berdasarkan nilai produksi kopi (dalam ribu ton). Metode yang digunakan adalah *clustering* hierarki dengan metode keterpautan *single linkage*, *complete linkage*, *average linkage* menggunakan jarak *Dynamic Time Warping* (DTW) dan *Euclidean*.

Dari hasil analisis diperoleh metode keterpautan yang terbaik serta jarak terbaik yaitu *average linkage* dengan jarak DTW. Metode *average linkage* dengan jarak DTW menunjukkan terbentuknya 3 *cluster*. Terbentuknya 3 *cluster* ini kemudian diuji menggunakan koefisien *silhouette* dan diperoleh hasilnya secara rata-rata sebesar 0.81 yang berarti *clustering* dengan 3 *cluster* dikategorikan ke dalam *Strong Classification*.

Tiga *cluster* yang dihasilkan adalah sebagai berikut.

- (1) *Cluster* pertama (2 provinsi) merupakan daerah dengan kategori produksi kopi "tinggi" setiap tahunnya di tahun 2008-2021, yakni Provinsi Sumatera Selatan dan Lampung.
- (2) *Cluster* kedua (4 provinsi) merupakan daerah dengan kategori produksi kopi "sedang" setiap tahunnya di tahun 2008-2021, yakni Provinsi Jawa Timur, Bengkulu, Aceh, dan Sumatera Utara.
- (3) *Cluster* ketiga (28 provinsi) merupakan daerah dengan kategori produksi kopi "rendah" setiap tahunnya di tahun 2008-2021, yakni Provinsi Kalimantan Tengah, Kalimantan Timur, Maluku Utara, Papua Barat, Kalimantan Utara, DKI Jakarta, Kep. Bangka Belitung, Kep. Riau, Gorontalo, DI Yogyakarta, Maluku, Kalimantan Selatan, Sulawesi Tenggara, Riau, Banten, Papua, Sulawesi Tengah, Kalimantan Barat, Sulawesi Utara, NTB, Sulawesi Barat, Jawa Tengah, NTT, Bali, Jambi, Jawa Barat, Sumatera Barat, dan Sulawesi Selatan.

UCAPAN TERIMA KASIH

Perlu kami sampaikan ucapan terima kasih kepada pihak-pihak yang terlibat dalam penyusunan jurnal ini, baik kepada dosen-dosen yang terlibat maupun teman-teman yang ikut terlibat dalam penyusunan ini.

DAFTAR PUSTAKA

- Adnyani, L. P., & Sihombing, P. R. (2021). ANALISIS CLUSTER TIME SERIES DALAM PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN NILAI PDRB. *I*(1).
- Annur, C. M. (2022, March 09). *Produksi Kopi Indonesia Naik Jadi 774,60 Ribu Ton pada 2021*. Retrieved March 27, 2023, from <https://databoks.katadata.co.id/datapublish/2022/03/09/produksi-kopi-indonesia-naik-jadi-77460-ribu-ton-pada-2021>
- Badan Pusat Statistik. *Produksi Tanaman Perkebunan (Ribu Ton)*. Retrieved March 27, 2023, from <https://www.bps.go.id/indicator/54/132/5/produksi-tanaman-perkebunan.html>
- Bhardwaj, A. (2020, May 26). *Silhouette Coefficient*. Retrieved April 16, 2023, from [towardsdatascience.com: https://towardsdatascience.com/silhouette-coefficient-validating-clustering-techniques-e976bb81d10c](https://towardsdatascience.com/silhouette-coefficient-validating-clustering-techniques-e976bb81d10c)
- Binty, T. (2019, July 06). *Hierarchical Clustering - Average Linkage with R*. Retrieved April 15, 2023, from [medium.com:](https://medium.com/)

- <https://medium.com/@story.of.stats/hierarchical-clustering-average-linkage-with-r-1310f93ed4f3>
- Briliant, E. H. (2019, June 21). *Analisis Cluster Hirarki Menggunakan R*. Retrieved April 16, 2023, from medium.com: <https://medium.com/@16611077/analisis-cluster-hirarki-menggunakan-r-ab8b23f8eb5a>
- Dani, A. T., Wahyuningsih, S., & Rizki, N. A. (2019). Penerapan Hierarchical Clustering Metode Agglomerative pada Data Runtun Waktu. *Jambura Journal of Mathematics*, 1(2), 64-78.
- Hartati, A. D. (2022, May 19). *Penerapan Analisis Cluster Untuk Mengelompokkan Kabupaten/Kota Berdasarkan Indikator Pembangunan Ekonomi di Provinsi Jawa*. Retrieved April 16, 2023, from https://rstudio-pubs-static.s3.amazonaws.com/905007_d5ab7298187a417081c2e04153d1ae35.html#koefisien-korelasi-cophenetic
- Hendry, J. (2012, January 30). *Dynamic Time Warping Distance*. Retrieved April 15, 2023, from scribd.com: <https://www.scribd.com/doc/100801150/Dynamic-Time-Warping-Distance#>
- Hidayat, A. (2023, January 02). *KLASTERISASI KASUS AKTIF COVID-19 DI JAWA BARAT BERDASARKAN KABUPATEN/KOTA*. Retrieved April 16, 2023, from rpubs.com: <https://rpubs.com/Anoe/project3>
- Humam, N. (n.d.). *Analisis Klaster (Hierarchical Clustering)*. Retrieved April 15, 2023, from rpubs.com: <https://rpubs.com/nadhifanhf/hierarchical-clustering>
- Kafritsas, N. (2021, September 14). *Time series classification using Dynamic Time*. Retrieved April 16, 2023, from towardsdatascience.com: <https://towardsdatascience.com/time-series-classification-using-dynamic-time-warping-61dcd9e143f6>
- Munthe, A. D. (2019). PENERAPAN CLUSTERING TIME SERIES UNTUK MENGGEROMBOLKAN PROVINSI DI INDONESIA BERDASARKAN NILAI PRODUKSI PADI. *JURNAL LITBANG SUKOWATI*, 11(2), 1-11.
- Pratiwi, A. M. (2022, May). *Analisis Cluster Hierarchy untuk Pengelompokan Kota/Kabupaten di Nusa Tenggara Timur berdasarkan Indikator Pendidikan Usia 15 Tahun Ke Atas*. Retrieved April 15, 2023, from rpubs.com: <https://rpubs.com/amelliamegadt/904898>
- Rizki, M. I., Taqqiyuddin, T. A., & Carelia, J. J. (2021). K-Medoids Clustering dengan Jarak Dynamic Time Warping dalam Mengelompokkan Provinsi di Indonesia Berdasarkan Kasus Aktif Covid-19. *PRISMA, PROSIDING SEMINAR NASIONAL MATEMATIKA*, 5, 685-692.
- Setiawan, D., & Zahra, A. (2023). Pengelompokan Kemiskinan di Indonesia Menggunakan Time Series Based Clustering. *INFERENSI*, 16(1), 83-90.
- Sobolewska, E. (n.d.). *Dynamic Time Warping (DTW) as a mean to cluster time series*. Retrieved April 15, 2023, from rpubs.com: <https://rpubs.com/esobolewska/dtw-time-series>
- Soleha, H. A., Nurmayanti, W. P., Hidayaturohman, U., Hirzi, R. H., & Septiani, A. (2022). Penerapan Clustering Time Series pada Pengelompokan Provinsi di Indonesia (Studi Kasus: Nilai Ekspor Non Migas di Indonesia Tahun 2016-2020). *J Statistika*, 15(2), 286-291.

Penerapan Clustering Time Series untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Jumlah Produksi Kopi Tahun 2008-2021

Ulinnuh, N., & Veriani, R. (2020). Analisis Cluster dalam Pengelompokan Provinsi di Indonesia Berdasarkan Variabel Penyakit Menular Menggunakan Metode Complete Linkage, Average Linkage, dan Ward. *Jurnal Nasional Informatika dan Teknologi Jaringan*, *V*(1).