

Analisis Perbandingan Pengelompokan Kota di Indonesia Berdasarkan Indikator Inflasi Tahun 2021 dengan Metode *Ward* dan *K-Means*

Lia Puspita Sari¹, Aris Fanani², Ahmad Hanif Asyhar³

^{1,2,3} Prodi Matematika, UIN Sunan Ampel Surabaya
Gunung Anyar, Kec. Gunung Anyar, Surabaya, Jawa Timur, 60294
Email: liapuspita011@gmail.com¹, arisfa@uinsby.ac.id², hanif@uinsby.ac.id³
*Korespondensi penulis : liapuspita011@gmail.com

Abstrak

Inflasi merupakan suatu kondisi perekonomian yang menunjukkan adanya kecenderungan kenaikan tingkat harga umum karena barang dan jasa yang ada di pasaran mempunyai jumlah dan jenis yang sangat beragam, sebagian besar dari harga-harga tersebut selalu meningkat dan mengakibatkan terjadinya inflasi. Perhitungan laju inflasi salah satunya adalah menggunakan Indeks Harga Konsumen (IHK), indeks ini disusun dari harga barang dan jasa yang dikonsumsi oleh masyarakat. Penelitian ini membahas tentang perbandingan pengelompokan kota di Indonesia berdasarkan indikator inflasi metode *Ward* dan *K-Means* dengan menggunakan 11 variabel kelompok pengeluaran pada IHK. Evaluasi metode dilakukan dengan membandingkan nilai rasio simpangan baku dari masing-masing metode. Berdasarkan hasil penelitian jumlah *cluster* yang dihasilkan adalah sebanyak 3 *cluster*, dengan nilai rasio simpangan baku yang diperoleh menggunakan metode *Ward* adalah 1,77, sedangkan dengan metode *K-Means* adalah sebesar 1,43. Dengan demikian hasil *cluster* yang bisa digunakan sebagai rujukan pada pengelompokan kota di Indonesia berdasarkan indikator inflasi adalah hasil analisis *cluster* menggunakan metode *K-Means*.

Kata Kunci: *Clustering*, Inflasi, *K-Means*, Simpangan Baku, *Ward*.

Abstract

Inflation is an economic condition that shows a tendency to increase in the general price level because the goods and services on the market have a very diverse number and types, most of these prices always increase and result in inflation. One way to calculate the inflation rate is to use the Consumer Price Index (CPI), this index is compiled from the prices of goods and services consumed by the public. This study discusses a comparison of the grouping of cities in Indonesia based on the *Ward* and *K-Means* method of inflation indicators using 11 expenditure group variables on the

CPI. Evaluation of the method is carried out by comparing the value of the standard deviation ratio of each method. Based on the research results, the number of clusters produced is 3 clusters, with the value of the standard deviation ratio obtained using the Ward method is 1.77, while the K-Means method is 1.43. Thus the cluster results that can be used as a reference for grouping cities in Indonesia based on inflation indicators are the results of cluster analysis using the K-Means method.

Keywords: Clustering, Inflation, K-Means, Standard Deviation, Ward.

1. Pendahuluan

Tingkat inflasi yang tinggi serta pertumbuhan perekonomian yang lambat adalah salah satu permasalahan ekonomi disuatu negara [1]. Ditengah masa pandemi, inflasi di Indonesia relatif terkendali sebesar 1,87% (yoy) dibandingkan dengan beberapa negara negara yang terus mengalami ketidakseimbangan penawaran permintaan dan krisis energi misalnya Singapura sebesar 3,8% (yoy), Euro Area sebesar 4,9% (yoy), dan Amerika Serikat sebesar 6,8% (yoy) pada November 2021 [2]. Inflasi adalah suatu kondisi perekonomian yang menunjukkan adanya kecenderungan kenaikan tingkat harga umum, karena barang dan jasa yang ada dipasaran mempunyai jumlah dan jenis yang sangat beragam dan sebagian besar dari harga-harga tersebut selalu meningkat dan mengakibatkan terjadinya inflasi [3]. Berhubung inflasi menyangkut harga sejumlah barang dan jasa, maka penghitungan inflasi tidak sesederhana menghitung kenaikan satu jenis barang dan jasa, penghitungan inflasi dapat menjadi sangat kompleks pada umumnya laju inflasi dihitung dengan menggunakan angka indeks [4].

Indeks Harga Konsumen (IHK) adalah indeks yang paling banyak digunakan dalam penghitungan inflasi. Indeks ini disusun dari harga barang dan jasa yang dikonsumsi oleh masyarakat. Jumlah barang dan jasa yang digunakan dalam penghitungan angka indeks tersebut berbeda antar negara dan antar waktu, bergantung pada pola konsumsi masyarakat akan barang dan jasa tersebut [5]. Tidak hanya negara saja yang mengalami inflasi tetapi pada lingkup yang lebih kecil yakni kota atau kabupaten juga mengalami inflasi.

Kota yang telah didata untuk dicatat inflasi dan IHK nya adalah 90 kota di Indonesia, dengan Jayapura adalah kota yang tertinggi inflasinya sebesar 1,91% [6]. Maka dari untuk dapat menekan dan menstabilkan inflasi, perlu mengetahui pengelompokan kota-kota tersebut sesuai IHK kelompok pengeluarannya. Lalu pemerintah bisa membuat sebuah kebijakan untuk menekan dan menstabilkan IHK sesuai kelompok-kelompoknya [7]. Dari latar belakang tersebut maka diperlukan analisis pengelompokan tingkat inflasi pada kota-kota di Indonesia dengan menggunakan metode *clustering*. *Clustering* adalah proses pengelompokan sekumpulan data ke dalam *cluster* yang memiliki kemiripan, dimana kesamaan dalam satu *cluster* dimaksimalkan, sedangkan kesamaan diluar *cluster* diminimalkan dengan perhitungan jarak [8].

Metode *clustering* dibagi menjadi dua yakni metode *hierarki* dan *non hierarki*. Metode *hierarki* merupakan teknik *clustering* yang membentuk tingkatan tertentu seperti struktur pohon dan hasilnya disajikan dalam bentuk dendogram, metode ini yakni *Single Linkage*, *Complete Linkage*, *Average Linkage*, *Ward*, dan *Centroid* [9]. Sedangkan metode *non hierarki*

prosedurnya dimulai dengan memilih sejumlah nilai *cluster* awal, kemudian objek pengamatan digabungkan ke dalam *cluster-cluster* tersebut, metode ini meliputi *Sequential Threshold*, *Parallel Threshold*, dan *Optimizing Partitioning*, salah satu metode *Optimizing Partitioning* yang sangat terkenal adalah *K-Means* [10].

Penelitian sebelumnya yang membahas tentang perbandingan hasil analisis *cluster* dengan menggunakan metode *Average Linkage* dan metode *Ward* oleh Imasdiani, dkk dengan studi kasus : kemiskinan di provinsi Kalimantan timur tahun 2018, diperoleh hasil rasio simpangan baku metode *Ward* lebih baik dari metode *Average Linkage* yaitu 2,681 [11]. Penelitian lain oleh Gurning dan Mustakim yaitu penerapan *K-Means* dan *K-Medoid* untuk pengelompokan data pasien covid-19, dengan hasil metode *K-Means* lebih optimal dibanding *K-Medoid* yang dibuktikan pada nilai DBI *K-Means* sebesar 0,139 pada percobaan $k = 4$ [12]. Penelitian selanjutnya oleh Novia Wulandari, dkk membahas tentang pengelompokan inflasi di Indonesia dengan metode *K-Means* dengan data tingkat inflasi kota di Indonesia tahun 2020 hingga 2022, diperoleh lima jumlah cluster yakni inflasi sangat tinggi, tinggi, sedang, rendah, dan sangat rendah [13]. Sedangkan penelitian ini adalah menganalisis perbandingan pengelompokan kota di Indonesia berdasarkan indikator inflasi dengan metode *Ward* dan *K-means*. Penelitian ini bertujuan untuk mencari metode terbaik yang bisa digunakan sebagai rujukan dalam mengelompokkan kota di Indonesia berdasarkan indikator inflasi tahun 2021.

2. Metode Penelitian

Jenis penelitian ini adalah penelitian kuantitatif, data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari situs Badan Pusat Statistik Indonesia (bps.go.id). Variabel yang digunakan adalah rata-rata bulanan dari 11 kelompok pengeluaran pada IHK yang disesuaikan dari 90 kota di Indonesia tahun 2021. Data IHK tersebut bisa dilihat di dalam Tabel 1.

Tabel 1. Definisi Variabel

No	Variabel	Indikator	Analisis Data
1	X_1	Perawatan pribadi jasa lainnya	Interval
2	X_2	Penyediaan makanan minuman restoran	Interval
3	X_3	Pendidikan	Interval
4	X_4	Rekreasi budaya olahraga	Interval
5	X_5	Informasi komunikasi jasa keuangan	Interval
6	X_6	Transportasi	Interval
7	X_7	Kesehatan	Interval
8	X_8	Perlengkapan peralatan pemeliharaan rutin rumah tangga	Interval
9	X_9	Perumahan air listrik bahan bakar rumah tangga	Interval
10	X_{10}	Pakaian alas kaki	Interval
11	X_{11}	Makanan minuman tembakau	Interval

Penelitian ini menggunakan metode *clustering*. *Clustering* merupakan metode untuk membagi uraian data menjadi beberapa *cluster* berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya. Sedangkan *cluster* adalah sekelompok objek-objek data yang sama antara satu dengan yang lain dalam *cluster* yang sama dan disamakan terhadap objek-objek yang berbeda *cluster*. Objek akan dikelompokkan ke dalam satu atau lebih *cluster* sehingga objek-objek yang berada dalam satu *cluster* akan mempunyai

kesamaan yang tinggi antara satu dengan yang lainnya [14]. *Clustering* termasuk ke dalam metode deskriptif dan pembelajaran tak terarah, dimana tidak ada pendefinisian kelas objek sebelumnya. Sehingga *clustering* bisa digunakan untuk menentukan label kelas bagi data yang belum diketahui kelasnya [15]. Langkah-langkah yang harus dilakukan dalam penelitian ini adalah pertama mendeskripsikan data untuk dapat mengetahui kespesifikan dari data setelah itu yang kedua melakukan standarisasi data menggunakan *decimal scaling*, ketiga melakukan *clustering* dengan metode *Hierarki Ward* dan *Non Hierarki Partitioning K-Means*, dan yang keempat adalah menemukan metode terbaik dengan menggunakan kriteria dua simpangan baku yakni S_w dan S_b .

Standarisasi Data

Standarisasi data dapat dilakukan jika pada variabel yang diteliti terdapat perbedaan ukuran satuan yang cukup besar. Perbedaan satuan ini akan menyebabkan perhitungan pada analisis *cluster* tidak benar-benar valid. Oleh karena itu dilakukan proses standarisasi data melalui transformasi pada data aktual sebelum melakukan analisis lebih lanjut. Transformasi dalam bentuk *decimal scaling* adalah seperti dibawah ini [16]:

$$v' = \frac{v}{10^J} \quad (1)$$

Keterangan

v' : Nilai baru setelah dilakukan *decimal scaling*.

v : Nilai atribut.

J : Bilangan bulat terkecil, sehingga $\max(|v'|) < 1$.

Hierarki Ward

Metode *Ward* termasuk metode varian yang biasa digunakan, metode ini menggabungkan *cluster*, jika total ketidaksamaan kuadrat dengan pusat *cluster* minimum disemua kemungkinan pilihan penggabungan [17]. Tujuan metode *Ward* ini adalah untuk memperoleh *cluster* yang memiliki varians internal yang sekecil mungkin. Metode *Ward* menggunakan ukuran *Sum of Square Error* (SSE) [18].

$$SSE = \sum_{i=1}^N (x_i - \bar{x})^2 \quad (2)$$

Keterangan

x_i : Vektor kolom yang entrinya nilai objek ke-I, dengan $i = 1, 2, 3, \dots, N$.

\bar{x} : Vektor kolom yang entrinya rata-rata nilai objek dalam *cluster*.

N : Banyaknya objek dalam *cluster* yang terbentuk.

Langkah-langkah dalam metode *Ward* adalah sebagai berikut [19] :

- a. Setiap objek atau pengamatan dianggap sebagai *cluster*, maka dari itu ditahap ini mempunyai N *cluster* dengan SSE bernilai nol.

- b. Setelah itu, membentuk *cluster* dengan memilih dua dari N *cluster* yang memiliki nilai SSE terkecil (N *cluster* berkurang satu pada setiap tahap $N - 1$).
- c. Membuat kombinasi dua pasang *cluster* baru yang terdiri dari satu *cluster* yang telah terbentuk dan memilih *cluster* lain, lalu hitung kembali SSE dari setiap pasang *cluster* baru dan memilih dua pasang *cluster* yang memiliki nilai SSE terkecil untuk digabungkan menjadi satu *cluster*.
- d. Ulangi langkah (3) sampai semua objek bergabung menjadi satu *cluster*.

Non Hierarki Partitioning K-Means

Partitioning non hirarki yang paling populer dan banyak digunakan pada *clusterisasi* data adalah metode *K-Means*. *K-Means* yaitu metode *clustering* berbasis jarak yang membagi data menjadi beberapa kelompok, algoritma ini sangat terkenal karena kemudahannya dan kemampuannya untuk mengelompokkan data yang besar dan data outlier dengan sangat cepat [20].

Metode *K-Means* mengelompokkan data dalam bentuk satu atau lebih *cluster*, data yang memiliki karakteristik yang sama dikelompokkan dalam satu *cluster*, dan data yang memiliki karakteristik yang berbeda dikelompokkan dengan *cluster* yang lain [21]. Karakteristik yang sama bisa diketahui dengan menerapkan pengukuran jarak, salah satu metode perhitungan jarak adalah *Euclidean Distance* [22]. *Euclidean Distance* adalah perhitungan jarak untuk 2 titik data x dan y dalam d dimensi data, rumus *Euclidean Distance* sebagai berikut.

$$d_{Euclidean\ Distance}(x, y) = \sqrt{\sum_{j=1}^d (x_j - y_j)^2} \quad (3)$$

Keterangan

- x, y : Dua titik diruang n -*Euclidean*.
 d : Vektor *Euclidean*.
 $x_j - y_j$: Ruang- n .

Tujuan data *clustering* ini untuk meminimalisasikan objek fungsi yang diset dalam proses *clustering*, yang umumnya berusaha meminimalisasikan variasi didalam suatu *cluster* dan memaksimalkan variasi antar *cluster* [21]. Langkah-langkah metode *K-Means* adalah :

- a. Tentukan banyaknya K *cluster* yang ingin dibentuk.
- b. Menemukan secara acak nilai *centroid* (pusat *cluster*) sejumlah K .
- c. Menghitung jarak tiap data terhadap masing-masing *centroid* *cluster* dengan menggunakan jarak *Eulidean*.
- d. Mengelompokkan data berdasarkan jarak terdekat.
- e. Menentukan *centroid* baru dengan menghitung rata-rata seluruh data pada pusat *cluster* yang sama.

$$C_{kj} = \frac{x_{1j} + x_{2j} + \dots + x_{aj}}{a}, \quad (4)$$

Keterangan :

C_{kj} : Pusat *cluster* ke-k variable ke j.

a : Banyaknya data pada *cluster* ke-k.

- f. Ulangi dari langkah (3) sampai anggota tiap *cluster* tidak ada yang berubah atau konvergen.

Metode Terbaik

Kualitas pengelompokan terbaik bisa diketahui dari kedua metode *cluster* yang dianalisis adalah dengan menggunakan kriteria dua simpangan baku, yaitu simpangan baku dalam *cluster* (S_w) dan simpangan baku antar *cluster* (S_b) dirumuskan dengan [23]:

$$S_w = \frac{1}{c} \sum_{k=1}^c S_k \quad (5)$$

Keterangan :

c : Jumlah *cluster*

K : Banyaknya kelompok yang terbentuk.

S_k : Simpangan baku kelompok ke-k.

$$S_b = \left[\frac{1}{c-1} \sum_{k=1}^c (\bar{X}_k - \bar{X})^2 \right]^{\frac{1}{2}} \quad (6)$$

dengan \bar{X}_k : Rataan kelompok k dan \bar{X} : Rataan keseluruhan kelompok.

Perbandingan metode *clustering*, dikatakan baik jika mempunyai nilai S_w yang minimum dan nilai S_b yang maksimum [24], artinya memiliki homogenitas yang tinggi. Metode yang dipilih adalah metode yang mempunyai nilai S_w/S_b terkecil.

3. Hasil dan Pembahasan

3.1 Deskripsi Data

Sebelum dilakukan proses perhitungan data, hal yang harus dilakukan adalah mengumpulkan data. Data yang digunakan dalam penelitian ini diperoleh dari Badan Pusat Statistik Indonesia (BPS) yaitu mengenai Indeks Harga Konsumen (IHK). Data 90 kota di Indonesia digunakan sebagai objek pengamatan, dengan karakteristik 11 kelompok pengeluaran pada IHK di Indonesia tahun 2021 secara menyeluruh disajikan pada Tabel 2.

Tabel 2. Analisis Statistik Deskriptif

VAR	Min	Max	Mean
X_1	103.0	116.3	109.0
X_2	94.32	120.43	106.31
X_3	96.49	107.74	102.40
X_4	97.5	115.6	106.7
X_5	100.3	140.6	109.5

X_6	93.66	117.15	103.32
X_7	88.82	106.14	99.91
X_8	97.67	128.74	107.03
X_9	97.05	123.50	107.02
X_{10}	97.48	138.26	107.68
X_{11}	105.1	125.3	113.4

3.2 Standardisasi Data

Data yang digunakan dalam penelitian ini mempunyai jangkauan jauh antar variabel, oleh karena itu harus dilakukan standardisasi data. Proses perhitungan standardisasi data dengan melakukan transformasi pada data asli menggunakan *decimal scaling* berdasarkan Persamaan (1). Adapun hasil standardisasi data tersebut dapat dilihat dalam Tabel 3.

Tabel 3. Hasil Standardisasi Data

No	Nama Kota	X_1	X_2	X_3	X_4	X_5	...	X_{11}
1	Kota Meulaboh	-0,365	-2,432	-3,066	-3,151	-0,628	⋮	1,153
2	Kota Banda Aceh	-0,279	0,068	-0,072	-1,014	1,141	⋮	0,008
3	Kota Lhokseumawe	0,970	0,806	-0,025	-0,661	-0,232	⋮	-0,987
4	Kota Sibolga	0,335	3,131	-0,897	1,428	-1,015	⋮	-0,367
5	Kota Pematang Siantar	0,518	-0,620	-0,964	-0,126	0,708	⋮	-1,717
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
86	Manokwari	-0,193	0,023	2,346	-0,053	0,422	⋮	2,657
87	Kota Sorong	1,218	-1,126	-0,736	-0,551	-0,251	⋮	0,253
88	Merauke	-0,351	-1,101	-0,969	-0,451	1,141	⋮	-0,119
89	Timika	1,900	-1,361	-2,869	-1,597	-1,408	⋮	-2,031
90	Kota Jayapura	-1,521	-0,695	-0,591	-0,736	-0,353	⋮	-1,668

3.3 Ward

Cluster yang terbentuk dari metode *hierarki Ward* adalah seperti Tabel 4, Dari 90 Kota di Indonesia terbagi menjadi tiga *cluster* yaitu *cluster* 1 memiliki 52 anggota, *cluster* 2 memiliki 34 anggota, dan *cluster* 3 memiliki 4 anggota.

Tabel 4. Hasil Clustering Ward

Cluster ke	Jumlah Anggota	Rata-Rata Variabel
1	52	105,545455
2	34	107,563636
3	4	111,909091
Jumlah Rata-Rata variabel		108,339394

Berdasarkan nilai rata-rata keseluruhan data yang diperoleh, dapat ditarik kesimpulan bahwa dengan metode *Ward*, *cluster* yang memiliki tingkat inflasi yang rendah adalah *cluster* 3 dengan nilai rata-rata variabel terbesar yaitu 111,909091, *cluster* yang memiliki tingkat inflasi yang sedang adalah *cluster* 2 dengan nilai rata-rata variabel 107,563636, dan *cluster* yang memiliki tingkat inflasi yang tinggi adalah *cluster* 1 dengan nilai rata-rata variabel terkecil yaitu 105,545455. Rincian data keanggotaan *cluster* metode *Ward* dapat dilihat pada Tabel 5.

Tabel 5. Hasil Keanggotaan Clustering Ward

Cluster	Nama Kota
1	Meulaboh, Banda Aceh, Pematang Siantar, Medan, Gunung Sitoli, Padang, Tembilahan, Pekanbaru, Dumai, Bungo, Palembang, Lubuklinggau, Bengkulu, Pangkal Pinang, Batam, Tanjung Pinang, DKI Jakarta, Cirebon, Tasikmalaya, Cilacap, Kudus, Surakarta, Banyuwangi, Sumenep, Kediri, Malang, Probolinggo, Madiun, Surabaya, Denpasar, Mataram, Bima, Waingapu, Maumere, Kupang, Pontianak, Singkawang, Sampit, Palangkaraya, Balikpapan, Tanjung Selor, Tarakan, Makassar, Palopo, Kendari, Bau-Bau, Gorontalo, Ternate, Sorong, Merauke, Timika, Jayapura.
2	Lhokseumawe, Sibolga, Bukittinggi, Jambi, Bandar Lampung, Metro, Tanjung Pandan, Bogor, Sukabumi, Bandung, Bekasi, Depok, Purwokerto, Semarang, Tegal, Yogyakarta, Jember, Tangerang, Cilegon, Serang, Ambon, Tual, Singaraja, Kota Baru, Tanjung, Banjarmasin, Samarindah, Manado, Kotamobagu, Luwuk, Palu, Watampone, Pare-pare, Mamuju.
3	Padangsidempuan, Sintang, Bulukumba, Manokwari.

3.4 K-Means

Cluster yang terbentuk dari metode *partitioning K-Means* adalah seperti Tabel 6, Dari 90 Kota di Indonesia terbagi menjadi tiga cluster yaitu cluster 1 memiliki 27 anggota, cluster 2 memiliki 9 anggota, dan cluster 3 memiliki 54 anggota.

Tabel 6. Hasil Clustering K-Means

Cluster ke	Jumlah Anggota	Rata-Rata Variabel
1	27	110,181818
2	9	107,409091
3	54	105,636364
Jumlah Rata-Rata variabel		107,742424

Berdasarkan nilai rata-rata keseluruhan data yang diperoleh, dapat ditarik kesimpulan bahwa dengan metode *K-Means*, cluster yang memiliki tingkat inflasi yang rendah adalah cluster 1 dengan nilai rata-rata variabel terbesar yaitu 110,181818, cluster yang memiliki tingkat inflasi yang sedang adalah cluster 2 dengan nilai rata-rata variabel 107,409091, dan cluster yang memiliki tingkat inflasi yang tinggi adalah cluster 3 dengan nilai rata-rata variabel terkecil yaitu 105,636364. Rincian data keanggotaan cluster metode *K-Means* dapat dilihat pada Tabel 7.

Tabel 7. Hasil Keanggotaan *Clustering K-Means*

<i>Cluster</i>	Nama Kota
1	Sibolga, Tembilahan, Bandar Lampung, Tanjung Pandan, Bogor, Sukabumi, Bandung, Bekasi, Depok, Semarang, Tegal, Yogyakarta, Jember, Surabaya, Tangerang, Cilegon, Singaraja, Kota Baru, Tanjung, Manado, Luwuk, Palu, Watampone, Pare-Pare, Gorontalo, Ambon, Kota Tual.
2	Padangsidempuan, Jambi, Metro, Serang, Sintang, Banjarmasin, Kotamobagu, Bulukumba, Manokwari .
3	Meulaboh, Banda Aceh, Lhokseumawe, Pematang Siantar, Medan, Gunungsitoli, Padang, Bukittinggi, Pekanbaru, Dumai, Bungo, Palembang, Lubuklinggau, Bengkulu, Pangkal Pinang, Batam, Tanjung Pinang, DKI Jakarta, Cirebon, Tasikmalaya, Cilacap, Purwokerto, Kudus, Surakarta, Banyuwangi, Sumenep, Kediri, Malang, Probolinggo, Madiun, Denpasar, Mataram, Bima, Waingapu, Maumere, Kupang, Pontianak, Singkawang, Sampit, Palangkaraya, Balikpapan, Samarindah, Tanjung Selor, Tarakan, Makassar, Palopo, Kendari, Bau-bau, Mamuju, Ternate, Sorong, Merauke, Timika, Jayapura.

3.5 Metode Terbaik

Pada Tabel 2 dan Tabel 4 dituliskan jumlah anggota *cluster* metode *Ward* dan *K-Means* memiliki perbedaan anggota yang cukup besar, perbedaan tersebut dikarenakan ada perbedaan dari rata-rata variabel. Pemilihan metode terbaik diantara dua metode dapat dilihat dari nilai simpangan baku yang disajikan dalam Tabel 8 berikut.

Tabel 8. Nilai Simpangan Baku

Simpangan Baku	<i>Ward</i>	<i>K-Means</i>
Sw	5,76298209	3,292814822
Sb	3,2519738	2,290987
Sw/Sb	1,77	1,43

Berdasarkan hasil tersebut menunjukkan bahwa nilai rasio simpangan baku (Sw/Sb) metode *K-Means* sebesar 1,43 memiliki nilai simpangan baku yang lebih kecil dari metode *Ward*. Sehingga kinerja metode yang paling baik dari kedua metode yang digunakan adalah metode *K-Means*. Hal itu sesuai dengan penelitian terdahulu yang sudah dibahas di pendahuluan yakni penerapan algoritma *K-Means* dan *K-Medoid* untuk pengelompokan data pasien covid-19, dengan hasil metode *K-Means* lebih optimal dibanding *K-Medoid* yang dibuktikan pada nilai DBI *K-Means* sebesar 0,139 pada percobaan $k = 4$ [12]. Maka dari itu *K-Means* adalah metode yang bisa dijadikan rujukan dalam penelitian ini karena algoritmanya memiliki kemampuan untuk mengelompokkan data yang besar dengan sangat cepat dan efisien.

4. Kesimpulan

Berdasarkan hasil analisis dan pengolahan data, dapat disimpulkan bahwa hasil perbandingan antara metode *Ward* dan *K-Means* pada nilai rasio simpangan baku pengelompokan kota di Indonesia berdasarkan indikator inflasi tahun 2021. Nilai rasio simpangan baku pada metode *K-Means* lebih kecil yakni 1,43 dibanding metode *Ward* yang sebesar 1,77, ini menunjukkan bahwa metode terbaik dalam pengelompokan kota di Indonesia berdasarkan indikator inflasi tahun 2021 adalah metode *K-Means*.

Daftar Pustaka

- [1] A. Salim, Fadilla, and A. Purnamasari, "Pengaruh Inflasi Terhadap Pertumbuhan Ekonomi Indonesia," *Ekonomica Sharia: Jurnal Pemikiran dan Pengembangan Ekonomi Syariah*, vol. 7, no. 1, pp. 17–28, 2021.
- [2] "Ditengah Peningkatan Inflasi Global, Laju Inflasi Indonesia Tahun 2021 Tetap Terkendali Rendah dan Stabil," *Kementerian Koordinator Bidang Perekonomian Republik Indonesia*, 2022.
- [3] S. Indriyani, "Analisis Pengaruh Inflasi dan Suku Bunga Terhadap Pertumbuhan Ekonomi di Indonesia Tahun 2005 – 2015," *Jurnal Manajemen Bisnis Krisnadwipayana*, vol. 4, no. 2, 2016, doi: 10.35137/jmbk.v4i2.37.
- [4] Suseno and S. Aisyah, "Inflasi Seri KebanKsentralan," *Seri Kebanksentralan*, no. 22, pp. iii–57, 2009.
- [5] G. A. D. Utari, R. Cristina, and S. Pambudi, "Inflasi di Indonesia: Karakteristik dan Pengendaliannya," *Bank Indonesia Institute*, vol. 23, no. 23, pp. 1–64, 2015.
- [6] BPS, *Perkembangan Indeks Harga Konsumen Desember 2021*. 2022.
- [7] M. Fuady dan J. Nugraha, "Implementasi Metode K-Means dan K-Medoids Untuk Mengelompokkan 82 Kota di Indonesia Berdasarkan Indeks Harga Konsumen," *Prosiding Seminar Nasional seri 7*, pp. 327–337, 2017.
- [8] J. Han and M. Kamber, "Data Mining : Concepts and Techniques (2nd edition)," pp. 1–6, 2006.
- [9] M. K. Rafsanjani, Z. A. Varzaneh, and N. E. Chukanlo, "A Survey of Hierarchical Clustering Algorithms," vol. 3, no. 3, pp. 229–240, 2012.
- [10] H. Prasetyo, W. Purnomo, Soenarnatalina, M. Adriani, and B. Wijanarko, "Penerapan Clustering Bootstrap dengan Metode K-Means," vol. 3, pp. 43–49, 2014.
- [11] Imasdiani, I. Purnamasari, and F. D. T. Amijaya, "Perbandingan Hasil Analisis Cluster dengan Menggunakan Metode Average Linkage dan Metode Ward (Studi Kasus : Kemiskinan Di Provinsi Kalimantan Timur Tahun 2018)," *Jurnal EKSPONENSIAL*, vol. 13, no. 1, pp. 9–18, 2022.
- [12] U. R. Gurning and Mustakim, "Penerapan Algoritma K-Means dan K-Medoid untuk Pengelompokan Data Pasien Covid-19," *Building of Informatics, Technology and Science (BITS)*, vol. 3, no. 1, p. 48–55, 2021, doi: 10.47065/bits.v3i1.1003.
- [13] N. Wulandari, nisa dienwati Nuris, and S. Anwar, "Penerapan Algoritma K-Means Clustering Pada Tingkat Inflasi Kota Di Indonesia," *Jurnal Riset Ilmu Akuntansi*, vol. 2, no. 2, pp. 15–34, 2023.
- [14] B. M. Metisen and H. L. Sari, "Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokan Penjualan Produk Pada Swalayan Fadhila," *Jurnal Media Infotama*, vol. 11, no. 2, pp. 110–118, 2015.

- [15] Parasian D. Silitonga and irene sri Morina, "Klusterisasi Pola Penyebaran Penyakit Pasien Berdasarkan Usia Pasien dengan Menggunakan K-Means Clustering," *Jurnal TIMES*, vol. VI, no. Vol 6, No 2 (2017), pp. 22–25, 2017.
- [16] P. P. Tjaya, Rino, and H. Wijaya, "Implementasi Metode Clustering K-Means Untuk Rekomendasi Pengadaan Stok Lampu di PT Global Lighting Indonesia," vol. 0577, no. I, 2021.
- [17] S. Landau and I. Chis Ster, "Cluster analysis: Overview," *International Encyclopedia of Education*, no. December, pp. 72–83, 2010, doi: 10.1016/B978-0-08-044894-7.01315-4.
- [18] A. N. Fathia, R. Rahmawati, and Tarno, "Analisis Klaster Kecamatan di Kabupaten Semarang Berdasarkan Potensi Desa Menggunakan Metode Ward Dan Single Linkage," *Jurnal Gaussian*, vol. 5, no. 4, pp. 801–810, 2016.
- [19] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, (5th editi. New York, 2001.
- [20] P. Berkhin, "Survey of Clustering Data Mining Techniques," *Technical Report, Accrue Software*, pp. 1–56, 2002.
- [21] Y. Agusta, "K-Means-Penerapan, Permasalahan dan Metode Terkait," *Jurnal Sistem dan Informatika*, vol. 3, no. Pebruari, pp. 47–60, 2007.
- [22] R. Hidayati, A. Zubair, A. Hidayat Pratama, and L. Indana, "Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering," *Techno.Com*, vol. 20, no. 2, pp. 186–197, 2021.
- [23] D. R. Ningrat, D. A. I. Maruddani, and T. Wuryandari, "Analisis Cluster dengan Algoritma K-Means dan Fuzzy C-Means Clustering Untuk Pengelompokan Data Obligasi Korporasi," *Jurnal Gaussian*, vol. 5, no. 4, pp. 641–650, 2016.
- [24] K. Arai and A. Ridho Barakbah, "Hierarchical K-means: an algorithm for centroids initialization for K-means," *Rep. Fac. Sci. Engrg. Reports of the Faculty of Science and Engineering*, vol. 36, no. 1, pp. 25–31, 2007.