

Comparison of Single Linkage, Complete Linkage, and Average Linkage Methods on Community Welfare Analysis in Cities and Regencies in East Java

Perbandingan Metode Single Linkage, Complete Linkage, dan Average Linkage pada Kesejahteraan Masyarakat di Jawa Timur

Yanuwar Reinaldi¹, Nurissaidah Ulinnuha², Tony Hartono³, Moh. Hafiyusholeh⁴

Abstract

Community welfare is one of the important points for a region and is also the essence of national development. The welfare of the people in Indonesia is fairly unequal, especially in East Java. To be able to map an area to the welfare of its people in East Java, one way that can be used is to use clustering. The hierarchical clustering method is one of the clustering methods for grouping data. In hierarchical clustering, single linkage, complete linkage, and average linkage methods are suitable methods for grouping data, which will compare the best method to use. The results of the calculation show that the average linkage method with three clusters is the best calculation with a silhouette index value of 0.6054, with the 1st cluster there are 23 regions, namely the city/district with the highest community welfare, the 2nd cluster there are 11 regions, namely cities/districts with moderate social welfare, and in the third cluster there are 4 regions, namely cities/districts with the lowest community welfare.

Keywords: Single linkage, Complete Linkage, Average Linkage

Abstrak

Kesejahteraan masyarakat merupakan salah satu poin penting bagi suatu daerah dan juga merupakan hakikat pembangunan nasional. Kesejahteraan masyarakat di Indonesia terbilang tidak merata, khususnya Jawa Timur. Untuk dapat memetakan suatu wilayah terhadap kesejahteraan masyarakat di Jawa Timur, salah satu cara yang dapat dipergunakan adalah menggunakan *clustering*. Metode *hierarchical clustering* merupakan salah satu metode *clustering* untuk mengelompokkan suatu data. Pada *hierarchical clustering*, metode *single linkage*, *complete linkage*, dan *average linkage* merupakan metode yang cocok untuk digunakan dalam pengelompokkan data, yang nanti akan dibandingkan metode yang terbaik untuk digunakan. Hasil dari perhitungan didapatkan bahwa metode *average linkage* dengan tiga klaster merupakan perhitungan yang paling baik dengan nilai *silhouette index* sebesar 0,6054, dengan klaster ke-1 terdapat 23 daerah yaitu kota/kabupaten dengan kesejahteraan masyarakat tertinggi, klaster ke-2 terdapat 11 daerah yaitu kota/kabupaten dengan kesejahteraan masyarakat sedang, dan klaster ke-3 terdapat 4 daerah yaitu kota/kabupaten dengan kesejahteraan masyarakat terendah.

Kata kunci: Single linkage, Complete Linkage, Average Linkage

*Program Studi Matematika, Universitas Islam Negeri Sunan Ampel

Email Address : h92218053@uinsby.ac.id¹, nuris.ulinnuha@uinsby.ac.id², tonyhartono@bps.go.id³, hafiyusholeh@uinsby.ac.id⁴



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)

1. Pendahuluan

Kesejahteraan masyarakat merupakan salah satu poin penting bagi suatu daerah dan juga merupakan hakikat pembangunan nasional. Kesejahteraan masyarakat sendiri dapat menggambarkan pada suatu daerah atau wilayah tersebut sejahtera atau tidak. Semakin tinggi suatu kesejahteraan maka kualitas tersebut juga semakin baik, dan juga sebaliknya [15]. Oleh karena itu, pemerintah sebaiknya berpikir dan mencari sebuah cara untuk meningkatkan kesejahteraan dengan cara meningkatkan pertumbuhan ekonomi yang tinggi di daerah tersebut dan juga proses pembangunan daerahnya [17].

Kesejahteraan masyarakat di Indonesia terbilang masih tidak merata antara provinsi satu dengan yang lainnya. Ini juga terjadi pada daerah-daerah yang ada di provinsi Jawa Timur. Hal ini terlihat dari suatu aspek antara satu daerah dengan lainnya yang berbeda. Aspek kesejahteraan suatu daerah dilihat dari beberapa aspek diantara lain aspek dalam pendidikan, IPM, kesehatan, PDRB dan lainnya. Menurut data Badan Pusat Statistik (BPS) Jawa Timur dalam buku indikator kesejahteraan masyarakat 2019-2020, kesenjangan kesejahteraan terjadi seperti pada variabel rata-rata lama sekolah. Nilai rata-rata lama sekolah di Jawa Timur yaitu sebesar 7,59 dengan daerah tertinggi di Madiun dengan nilai 11,13 dan terendah di Sampang dengan nilai 4,55. Untuk aspek lainnya yaitu persentase kemiskinan dengan nilai persentase kemiskinan di Jawa Timur sebesar 10,20 persen, dengan tingkat persentase terkecil yaitu di kota Batu dan tertinggi pada tiga daerah di Madura yaitu, Sampang, Sumenep, dan Bangkalan [5]. Dilihat dari data setiap aspek tersebut terlihat suatu kesenjangan kesejahteraan diantara suatu kabupaten/kota di Jawa Timur dengan lainnya. Oleh karena itu diperlukan suatu cara untuk dapat memetakan kesejahteraan masyarakat tersebut.

Terdapat beberapa cara yang dapat digunakan untuk dilakukan untuk memetakan suatu wilayah terhadap kesejahteraan masyarakatnya. *Clustering* merupakan suatu klasifikasi dari metode data mining untuk membagi grup dari suatu rangkaian data dengan titik kesamaannya yang mirip. *Clustering* terdapat beberapa macam yaitu *hierarchial* dan *non-hierarchial clustering* [9]. *Hierarchial clustering* merupakan salah satu jenis analisis kluster yang bertujuan dalam mengelompokkan suatu grup dengan melihat kesamaan karakteristiknya [14]. Dalam *hierarchial* dibagi menjadi *agglomerative* dan *divisive*. Pada *agglomerative* terbagi menjadi beberapa metode yaitu *single linkage*, *complete linkage*, *average linkage*, *ward method*, *WPGMA method*, *median method*, dan *centrid method* [6]. Metode *single linkage*, *complete linkage*, dan *average linkage* sendiri merupakan salah satu cara yang dapat digunakan dalam permasalahan kesejahteraan masyarakat pada kota di Jawa Timur dengan mengelompokkan kabupaten/kotanya.

Alwi dan Hasrul [2] melakukan penelitian tentang *clustering* kabupaten/kota di Sulawesi Selatan menggunakan *hierarchial clustering* dengan metode *average linkage* dan hasil yang didapatkan adalah pengelompokkan kabupaten/kota di provinsi Sulawesi Selatan dibagi menjadi tiga kluster, dengan kluster pertama terdapat 21 kabupaten/kota, untuk kluster kedua terdapat 2 kabupaten/kota, dan kluster ketiga hanya terdapat 1 kabupaten/kota. Ningsih, Wahyuningsih dan Nasution [13] juga melakukan penelitian tentang pengelompokan produksi palawija menggunakan *hierarchial clustering* dengan membandingkan antara *complete linkage* dan *average linkage*. Hasil penelitian adalah metode *average linkage* lebih baik, dikarenakan nilai simpangan baku *average linkage* 0,056 lebih kecil daripada *complete linkage*. Mu'afa dan Ulinuha [11] dengan penelitiannya yaitu pengelompokan kecamatan berdasarkan nilai variabel jenis ternak dan metode yang digunakan adalah tiga metode yang terdiri dari, *single linkage*, *complete linkage*, dan *average linkage* yang nantinya akan dibandingkan untuk metode terbaik. Hasil dari penelitian tersebut adalah metode *complete linkage* merupakan yang paling baik,

dikarenakan dilihat dari nilai rasio simpangan baku dari metode *complete linkage* bernilai 0,222 lebih kecil dari yang lainnya.

Pada penelitian ini, metode yang digunakan tidak hanya *average linkage* saja tetapi juga *single linkage* dan *complete linkage*. Setiap metode dalam *clustering* memiliki kelebihan dan kelemahan masing-masing, akan tetapi secara umum analisis kluster memiliki kelebihan yaitu dapat mengelompokkan data dalam jumlah banyak dan variabel yang banyak, sedangkan kelemahannya yaitu pengelompokkan bersifat subjektifitas, dan data yang terlalu berbeda akan sulit dikelompokkan. Oleh karena itu, untuk mengetahui metode yang unggul, dilakukan perbandingan dengan mengamati hasil dari setiap validasi *clustering*. Untuk hasil *clustering* diharapkan kepada pemerintah, khususnya pemerintah provinsi Jawa Timur untuk dijadikan bahan pertimbangan kedepannya untuk menyamaratakan kesejahteraan masyarakat di seluruh kabupaten/kota di Jawa Timur supaya tidak ada kesenjangan antara satu kabupaten/kota dengan yang lain.

2. Tinjauan Pustaka

2.1 Kesejahteraan Masyarakat

Kesejahteraan adalah suatu tolok ukur apakah sudah dapat memenuhi kondisi sandang, pangan, dan papannya. Tidak hanya dalam masalah kebutuhan hidup, kesejahteraan juga mencakup tentang kualitas hidup seperti mendapatkan pendidikan dan kesehatan yang memadai. Kesejahteraan dilaksanakan supaya masyarakat mengembangkan dirinya menjadi lebih baik dan dapat hidup yang layak [7]. Semakin tinggi kesejahteraan maka semakin baik pula kehidupan yang didapatkan.

Jadi kesejahteraan masyarakat adalah suatu acuan untuk suatu terpenuhinya kondisi di dalam masyarakat atau daerah tersebut terhadap kebutuhan dasar yang dapat dicerminkan dari tercukupinya kebutuhan hidupnya atau setiap individu dapat memaksimalkan nilai guna atau utilitasnya pada saat berada di tingkat batas anggaran tertentu [18].

Di Jawa Timur, kesejahteraan masyarakatnya dapat dibidang masih belum merata. Hal tersebut dapat dilihat dari data menurut Badan Pusat Statistik (BPS) Jawa Timur yang menunjukkan bahwa nilai suatu aspek kesejahteraan antara daerah satu dengan lainnya terjadi kesenjangan. Seperti contohnya pada aspek pendidikan, kemiskinan, IPM, dan lain-lain. Jadi diperlukan suatu langkah atau terobosan baru oleh pemerintah untuk menyamaratakan hal ini.

2.2 Hierarchical Clustering

Hierarchical clustering adalah salah satu cara mengelompokkan suatu data dari dua dokumen atau lebih dengan kesamaan yang dekat. Setelah itu dilanjutkan pada dokumen lainnya dengan kesamaan kedua yang dekat. Dalam *hierachical sendiri* terbagi menjadi dua metode klasifikasi yaitu *agglomerative* dan *divisive*. Dalam *agglomerative* terdapat beberapa metode yang digunakan yaitu sebagai berikut [8]:

Pertama adalah *single linkage* merupakan salah satu metode yang ada pada *hierarchical clustering* yaitu mengelompokkan dokumen atau objek dengan jarak terdekat antar dokumen tersebut. Langkah-langkah dalam menggunakan *single linkage* yaitu, langkah pertama dimulai dari K kluster, yaitu dari setiap kluster mempunyai sebuah matriks yang baris dan kolomnya simetrik dari jarak D dengan tipe $K \times K$. Langkah selanjutnya yaitu mencari pasangan kluster dari matriks jarak yang terdekat, dimisalkan dua kluster yaitu A dan B . Setelah mencari pasangan kluster, langkah selanjutnya menggabungkan antara kluster A dan B dengan bentuk

klaster (AB). Ubah setiap entri pada matriks yaitu dengan menghapus kolom dan baris yang bertepatan dan juga memberikan tambahan kolom dan baris dari klaster yang tersisa dan jarak antara klaster (AB). Terakhir yaitu mengulangi langkah pada kedua dan ketiga sebanyak $(K - 1)$ kali [16].

Kedua adalah *complete linkage*, yaitu suatu metode *clustering* yang mengelompokkan suatu objek atau dokumen dari jarak terjauh. Tahapan pada *complete linkage* hampir dengan tahapan pada *single linkage*, hanya berbeda pada pencarian pasangan yaitu jarak yang jauh atau maksimum [20].

Ketiga adalah *average linkage*, yaitu metode *clustering* yang mengelompokkan suatu dokumen atau objek dengan jarak rata-rata. Jarak tersebut dihitung disetiap klaster dengan persamaan berikut :

$$d_{(UV)W} = \frac{\sum_i \sum_k d_{ik}}{N_{(UV)}N_W}$$

Objek k berada dalam klaster W dan d_i adalah jarak diantara objek i . Untuk N_w dan $N_{(UV)}$ sendiri secara berturut-turut adalah jumlah objek di dalam klaster W dan UV [12].

2.3 Silhouette Index

Silhouette index adalah satu cara yang dapat digunakan untuk mengetahui kekuatan klaster dan melihat kualitasnya. Nilai *silhouette index* yang baik dan bagus untuk dipakai adalah yang mendekati angka 1, semakin mendekati maka jumlah klaster itulah yang baik digunakan [3]. Tahapan dalam menghitung *nilai silhouette index* yaitu, pertama menghitung nilai dari rata-rata jarak dari suatu objek dimisalkan dengan i dengan semua dari objek lainnya yang berada di dalam satu klaster dengan persamaan :

$$a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i, j)$$

Langkah selanjutnya adalah menghitung nilai dari rata-rata jarak dari suatu objek i dengan semua dari objek lainnya yang berada pada klaster lainnya, lalu ambil nilai paling maksimum dan minimum. Setelah menghitung nilai dari rata-rata objek, tahapan selanjutnya adalah menghitung nilai dari rumus *silhouette index* berikut :

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

Tahapan terakhir adalah membandingkan hasil dari nilai *silhouette index*. Nilai *silhouette index* berada diantara -1 hingga 1. Semakin mendekati angka 1, maka data semakin baik [4].

2.4 KMO dan Barlett's Test

KMO (*Kaiser-Meyer-Olkin*) adalah indeks dimana untuk membandingkan dua hal yaitu besarnya suatu koefisien dengan besarnya suatu koefisien korelasi yang sedang diamati. Hasil uji KMO yang baik adalah jika nilai lebih besar dari 0,5 atau juga dapat dilihat pada nilai *Measuring of Sample* (MSA) setiap variabel. Dari hasil nilai tersebut maka analisis faktor bisa diproses. Jadi variabel yang layak digunakan pada uji KMO yaitu jika hasil pada uji KMO atau MSA lebih dari 0,5, yang mengartikan bahwa terdapat hubungan antar variabel satu dengan yang lain dan menerangkan variabel lainnya [10]. *Barlett's Test of Sphericity* merupakan suatu

tes yang digunakan untuk menguji interdependensi diantara variabel yang digunakan sebagai indikator. Kevalidan nilai bartlett's test yaitu jika nilai kurang dari 0,5 sehingga analisis faktor dapat dilanjutkan [19].

3. Metode Penelitian

Penelitian ini menggunakan metode *clustering* yang terdiri dari, *single linkage*, *complete linkage*, dan *average linkage* yang akan digunakan untuk mengelompokkan wilayah berdasarkan kesejahteraan masyarakatnya. Data pada penelitian ini didapatkan dari website Badan Pusat Statistik tahun 2020 pada 38 kota dan kabupaten di Provinsi Jawa Timur. Terdapat lima variabel yang digunakan pada penelitian pengelompokkan kesejahteraan masyarakat ini, yaitu data angka harapan hidup (X1), harapan lama sekolah (X2), IPM (X3), presentase penduduk miskin (X4), dan rata-rata lama sekolah (X5).

Langkah-langkah perhitungan dalam pada penelitian ini, yaitu sebagai berikut :

1. Mencari data yang dibutuhkan untuk dijadikan variabel sebagai acuan pada perhitungan *clustering*.
2. Mengumpulkan data-data yang didapatkan untuk dikelompokkan.
3. Melakukan uji KMO dan bartlett's test yang digunakan untuk mengetahui apakah data tersebut layak uji.
4. Melakukan pengolahan data pengklasteran dengan metode *single linkage*, *complete linkage*, dan *average linkage* hingga diperoleh klaster.
5. Melakukan uji *silhouette index* ntuk mendapatkan klaster dan metode terbaik.
6. Menentukan urutan klaster dari yang tertinggi hingga terendah dengan cara melihat nilai *centroid* terbesar dengan klaster tertinggi hingga *centroid* terkecil sebagai klaster rendah.

4. Hasil dan Pembahasan

Data pada penelitian ini merupakan data dari 38 kabupaten/kota di Jawa Timur dengan lima variabel yaitu data angka harapan hidup, harapan lama sekolah, IPM, presentase penduduk miskin, dan rata-rata lama sekolah pada tahun 2020. Sebelum dilakukan analisis klaster, maka terlebih dahulu dilakukan analisis secara deskriptif.

Tabel 1. Analisis Statistik Deskriptif

Var	Min	Maks	Mean	Std Deviasi
X1	66,74	74,18	71,64	1,97
X2	11,60	15,51	13,32	0,90
X3	62,70	82,23	71,87	4,98
X4	3,89	22,78	11,02	4,51
X5	4,85	11,14	7,94	1,51

Dilihat pada tabel 4.1, didapatkan beberapa nilai yaitu nilai minimal, maksimum, rata-rata (*mean*), dan standar deviasi dari setiap variabel. Pada variabel X1, rata-rata angka harapan hidup di kabupaten/kota di Jawa Timur adalah 71,64 dengan nilai minimal 66,74 pada kabupaten Bondowoso dan nilai maksimum 74,18 pada kota Surabaya. Pada variabel X2 menunjukkan bahwa harapan lama sekolah pada kabupaten/kota di Jawa Timur adalah 13,32, dengan nilai minimum 11,60 di Bangkalan dan maksimum 15,51 di kota Malang. Variabel X3 menunjukkan

Jurnal Matematika, Statistika & Komputasi
Yanuwar Reinaldi, Nurissaidah Ulinnuha, Tony Hartono, Moh. Hafiyusholeh

rata-rata IPM pada kabupaten/kota di Jawa Timur adalah 71,87, dengan nilai minimum 62,70 di Sampang dan maksimum 82,23 di kota Surabaya. Pada variabel X4 menunjukkan bahwa rata-rata persentase penduduk miskin pada kabupaten/kota di Jawa Timur sebesar 11,02, dengan nilai minimum 3,89 di kota Batu dan maksimum 22,78 di Sampang. Pada variabel X5 menunjukkan rata-rata lama sekolah di Jawa Timur adalah sebesar 7,94, dengan nilai minimum 4,85 di Sampang dan maksimum 11,14 di kota Madiun.

Sebelum dilakukan uji asumsi, dilakukan sebuah standarisasi data. Standarisasi dilakukan dengan tujuan data yang inputan yang tidak sama dijadikan data yang standar menggunakan suatu cara tertentu. Setelah data distandarisasi, jika suatu nilai tidak berada diantara angka $\pm 2,5$ maka nilai tersebut dikatakan *outlier*. Pada perhitungan didapatkan nilai diantara 2 hingga -2, dimana tidak ada data yang *outlier*. Setelah dilakukan standarisasi data, maka selanjutnya adalah uji asumsi atau analisis data.

4.1 Uji KMO dan Barlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,778
Bartlett's Test of Sphericity	Approx Chi Square	222,393
	Df	10
	Sig.	0,000

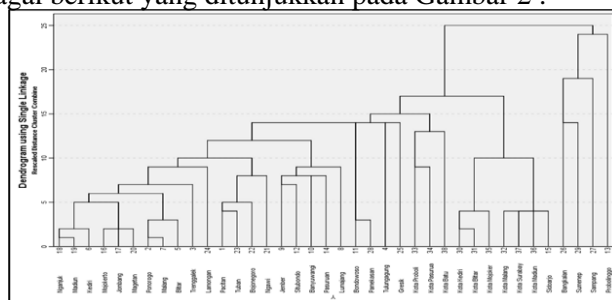
Tabel 2. Hasil Uji KMO dan *Barlett's Test*

Pada perhitungan uji KMO dan bartlett's test didapatkan hasil dari lima variabel yaitu sebagai berikut yang ditunjukkan pada Tabel 2.

Pada uji KMO dan *bartlett's test* sendiri berfungsi untuk mengetahui apakah variabel yang digunakan dikatakan layak atau tidak. Variabel yang layak yaitu, jika syarat pertama hasil dari uji KMO bernilai lebih dari 0,5 dan untuk *bartlett's test* bernilai kurang dari 0,5. Dari hasil perhitungan terhadap variabel kesejahteraan masyarakat, didapatkan hasil uji KMO yaitu sebesar 0,778 dengan variabel angka harapan hidup sebesar 0,828, variabel harapan lama sekolah sebesar 0,754, variabel IPM sebesar 0,691, variabel persentase penduduk miskin sebesar 0,968, dan variabel rata lama sekolah sebesar 0,755. Sedangkan pada bartlett's test bernilai sebesar 0. Dari hasil tersebut diketahui bahwa hasil uji KMO dan *bartlett's test* memenuhi persyaratan dan dapat dikatakan bahwa analisis faktor dapat dilanjutkan.

4.2 Analisis Clustering

Perhitungan menggunakan metode *single linkage* didapatkan hasil dalam bentuk dendrogram yaitu sebagai berikut yang ditunjukkan pada Gambar 2 :



Gambar 2. Hasil clustering *single linkage*

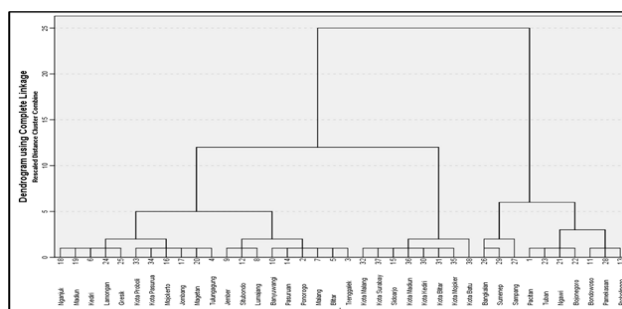
Jurnal Matematika, Statistika & Komputasi
Yanuwar Reinaldi, Nurissaidah Ulinnuha, Tony Hartono, Moh. Hafiyusholeh

Klaster 1 : Kota Batu, Kota Surabaya, Kota Madiun, Kota Mojokerto, Kota Pasuruan, Kota Probolinggo, Kota Malang, Kota Blitar, Kota Kediri, Pamekasan, Gresik, Lamongan, Tuban, Bojonegoro, Ngawi, Magetan, Madiun Nganjuk, Jombang, Mojokerto, Sidoarjo, Pasuruan, Situbondo, Bondowoso, Banyuwangi, Jember, Lumajang, Malang, Kediri, Blitar, Tulungagung, Trenggalek, Ponorogo, dan Pacitan.

Klaster 2 : Sampang, Sumenep, dan Bangkalan.

Klaster 3 : Probolinggo.

Perhitungan metode *complete linkage* didapatkan hasil dalam bentuk *dendrogram* yaitu sebagai berikut yang ditunjukkan pada Gambar 3 :



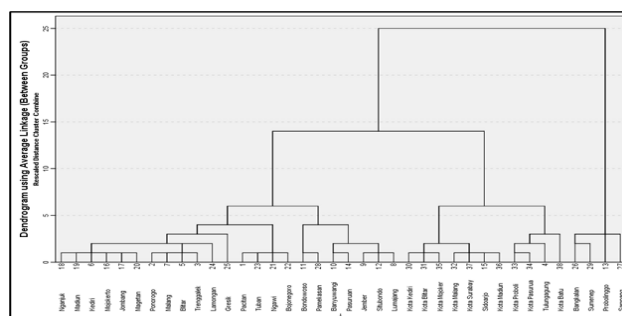
Gambar 3. Hasil clustering *complete linkage*

Klaster 1 : Kota Mojokerto, Kota Malang, Kota Blitar, Kota Kediri, Sidoarjo, Kota Batu, Kota Surabaya, dan Kota Madiun.

Klaster 2 : Kota Pasuruan, Kota Probolinggo, Gresik, Lamongan, Magetan, Madiun Nganjuk, Jombang, Mojokerto, Pasuruan, Situbondo, Banyuwangi, Jember, Lumajang, Malang, Kediri, Blitar, Tulungagung, Trenggalek, dan Ponorogo.

Klaster 3 : Sampang, Sumenep, Bangkalan, Pamekasan, Tuban, Bojonegoro, Ngawi, Probolinggo, Bondowoso, dan Pacitan.

Perhitungan metode *average linkage* didapatkan hasil dalam bentuk *dendrogram* yaitu sebagai berikut yang ditunjukkan pada Gambar 4 :



Gambar 4. Hasil clustering *average linkage*

Klaster 1 : Tulungagung, Sidoarjo, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota batu.

Klaster 2 : Pacitan, Ponorogo, Trenggalek, Blitar, Kediri, Malang, Lumajang, Jember, Banyuwangi, Bondowoso, Situbondo, Pasuruan, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Gresik, dan Pamekasan.

Klaster 3 : Probolinggo, Bangkalan, Sampang, dan Sumenep.

4.3 Uji Silhouette Index

Pada uji *silhouette index* setiap hasil dari setiap *clustering* pada setiap metode akan digunakan untuk mengetahui berapa klaster yang terbaik untuk digunakan. Setelah perhitungan yang dilakukan, berikut adalah hasil nilai *silhouette index* :

Tabel 3. Hasil uji *silhouette index*

Metode/Nilai Silhouette	2 klaster	3 klaster	4 klaster
Single Linkage	0,5571	0,4043	0,4386
Complete Linkage	0,5618	0,5947	0,5140
Average Linkage	0,5771	0,6054	0,5566

Dari hasil tersebut diketahui bahwa metode *clustering* terbaik yaitu menggunakan metode *average linkage* dengan 3 klaster, dikarenakan memiliki nilai tertinggi yaitu 0,6054 dari hasil yang lainnya. Menurut Rousseeuw dan Kaufman menjelaskan bahwa struktur pada nilai *silhouette index* dibagi empat yaitu *strong* dimana nilai *silhouette* berada diantara 0,7 hingga 1, kedua *medium* dimana nilai *silhouette* berada diantara 0,5 hingga 0,7, ketiga *weak* dimana nilai *silhouette* berada diantara 0,25 hingga 0,5, dan *no structure* dimana nilai *silhouette* kurang dari 0,25. Hasil perhitungan nilai *silhouette index* pada kesejahteraan masyarakat pada kota/kabupaten di Provinsi Jawa Timur didapatkan nilai *silhouette* yaitu 0.6054, dimana nilai tersebut termasuk dalam struktur *medium*

4.4 Interpretasi Klaster

Setelah didapatkan semua hasil dari nilai *silhouette index* bahwa metode *average linkage* dengan 3 klaster terbaik, maka langkah selanjutnya yaitu menentukan urutan klaster dari setiap daerah dari yang tertinggi hingga terendah. Perhitungan urutan klaster yaitu menggunakan nilai *centroid* dari setiap klaster. pada perhitungan ini diambil dari variabel-variabel yang digunakan dan didapatkan hasil yaitu sebagai berikut yang ditunjukkan pada Tabel 2:

Tabel 4. Hasil nilai centroid

Klaster/ Variabel	X1	X2	X3	X4	X5
Ke – 1	73,08	14,35	77,92	6,09	9,80
Ke – 2	71,73	13,00	70,20	11,72	7,44
Ke – 3	69,20	12,38	64,82	20,53	5,65

Diketahui bahwa setiap variabel yang digunakan memiliki arah yang berbeda, dilihat dari nilai centroid yang didapatkan. Pada variabel angka harapan hidup, semakin tinggi nilai maka kota/kabupaten tersebut memiliki angka harapan hidup yang baik. Pada variabel harapan lama sekolah, semakin tinggi nilai maka kota/kabupaten tersebut memiliki harapan lama sekolah yang baik. Pada variabel IPM, semakin tinggi nilai maka kota/kabupaten tersebut memiliki IPM yang baik. Pada variabel persentase penduduk miskin, semakin rendah nilai maka kota/kabupaten tersebut memiliki persentase penduduk miskin yang baik. Pada variabel rata-rata

lama sekolah, semakin tinggi nilai maka kota/kabupaten tersebut memiliki rata-rata lama sekolah yang baik.

Klaster 1 memiliki nilai *centroid* yang tinggi. Ini mengartikan bahwa kota/kabupaten pada klaster ke-1 tingkat kesejahteraannya tinggi. Kota/kabupaten pada klaster ke-1, Tulungagung, Sidoarjo, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota batu. Selanjutnya yaitu klaster ke-2 yaitu klaster dengan nilai tengah, ini mengartikan bahwa kota/kabupaten pada klaster ke-2 tingkat kesejahteraannya sedang. Kota/kabupaten pada klaster ke-2 yaitu, Pacitan, Ponorogo, Trenggalek, Blitar, Kediri, Malang, Lumajang, Jember, Banyuwangi, Bondowoso, Situbondo, Pasuruan, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Gresik, dan Pamekasan. Sedangkan pada klaster ke-3 yaitu klaster dengan nilai terendah, ini mengartikan bahwa kota/kabupaten pada klaster ke-2 tingkat kesejahteraannya rendah. Kota/kabupaten pada klaster ke-3 yaitu, Probolinggo, Bangkalan, Sampang, dan Sumenep. Dari hasil didapatkan dan dibandingkan dengan penelitian sebelumnya yang dilakukan Zainal Abidin [20] tentang *clustering* indikator kemiskinan di Jawa Timur, hasil pengelompokan kabupaten/kota yaitu hampir sama, dengan metode yang digunakan peneliti yaitu menggunakan *hierachihcal clustering* dengan pembagian klaster yaitu empat klaster.

5. Kesimpulan

Setelah dilakukan perhitungan dengan menggunakan tiga metode, yaitu *single linkage*, *complete linkage*, dan *average linkage*, didapatkan bahwa pada perhitungan analisis kesejahteraan masyarakat metode *average linkage* dengan 3 klaster merupakan yang terbaik dengan nilai *silhouette index* yaitu 0,6054. Kota/kabupaten dengan kesejahteraan tertinggi yaitu, Tulungagung, Sidoarjo, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, dan Kota batu. Selanjutnya kota/kabupaten dengan kesejahteraan sedang yaitu, Pacitan, Ponorogo, Trenggalek, Blitar, Kediri, Malang, Lumajang, Jember, Banyuwangi, Bondowoso, Situbondo, Pasuruan, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Ngawi, Bojonegoro, Tuban, Lamongan, Gresik, dan Pamekasan. Terakhir yaitu kota/kabupaten dengan kesejahteraan terendah yaitu Probolinggo, Bangkalan, Sampang, dan Sumenep.

Daftar Pustaka

- [1]. Abidin, Zainal., 2017. Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Indikator Kemiskinan dengan Menggunakan Analisis Cluster Hierarki. *Skripsi*, Fakultas Vokasi, Departemen Statistika Bisnis, Institut Teknologi Sepuluh November, Surabaya.
- [2]. Alwi, Wahidah, & Muh. Hasrul., 2018. Analisis Klaster untuk Pengelompokan Kabupaten/Kota di Propinsi Sulawesi Selatan Berdasarkan Indikator Kesejahteraan Masyarakat. *Jurnal MSA*, Vol. 6, No. 1, pp. 35-42.
- [3]. Anggara, Mario., Herry Sujiani, & Helfi Nasution., 2016. Pemilihan Distance Measure Pada K-Means Clustering Untuk Pengelompokan Member di Alvaro Fitness. *Jurnal Sistem dan Teknologi Informasi*, Vol. 1, No. 1, pp. 1-6.
- [4]. Asmiatun, Siti., Nur Wakhidah, & Astrid Novita Putri., 2020. Penerapan Metode K-Medoids Untuk Pengelompokan Kondisi Jalan di Kota Semarang. *Jurnal Teknik Informatika dan Sistem Informasi*, Vol. 6, No. 2, pp. 171-180.
- [5]. Badan Pusat Statistik Jawa Timur., 2019. *Indikator Kesejahteraan Rakyat Provinsi Jawa Timur*. Badan Pusat Statistik Jawa Timur., Surabaya.

- [6]. Dani, Andrea Tri Rian., Sri Wahyuningsih, & Nanda Arista Rizki., 2019. Penerapan Hierarchical Clustering Metode Agglomerative Pada Data Runtun Waktu. *Jambura Journal of Mathematics*, Vol. 1, No. 2, pp. 64-78.
- [7]. Elvina & Musdhalifah., 2019. Peningkatan Kesejahteraan Masyarakat melalui Partisipasi dan Implementasi Kebijakan dengan Efektivitas Pembangunan Program Dana Desa sebagai Variabel Intervening. *Jurnal Sosial Humaniora dan Pendidikan*, Vol. 3, No. 1, pp. 1-9.
- [8]. Goreti, Maria., Yuki Novia N, & Sri Wahyuningsih. 2016. Perbandingan Hasil Analisis Cluster dengan Menggunakan Metode Single Linkage dan Metode C-Means. *Jurnal EKSPONENSIAL*, Vol. 7, No. 1, pp. 9-16.
- [9]. Gustientiedinda, M. Hasmil Adiya, & Yenny Desnelita., 2019. Penerapan Algoritma K-Means untuk Clustering Data Obat-Obatan pada RSUD Pekanbaru. *Jurnal Nasional Teknologi dan Sistem Informasi*, Vol. 05, No. 1, pp. 17-24.
- [10]. Lintong, Rocky., Yohanes A. R. Langi, & Charles E. Mongi. 2019. Penerapan Analisis Faktor Terhadap Kualitas dan Kepuasan Pelayanan Pasien Rumah Sakit TK-II R.W Mongisidi. *Jurnal Matematika dan Aplikasi*, Vol. 9, No. 1, pp. 24-30.
- [11]. Mu'afa, Sulthan Fikri, & Nurissaidah Ulinuha., 2019. Perbandingan Metode Single Linkage, Complete Linkage, dan Average Linkage dalam Pengelompokan Kecamatan Berdasarkan Variabel Jenis Ternak Kabupaten Sidoarjo. *Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, Vol.4, No. 2.
- [12]. Nasifah, Qonitatin, & Novita Eka Chandra., 2017. Analisis Cluster Average Linkage Berdasarkan Faktor-Faktor Kemiskinan di Provinsi Jawa Timur. *Math Journal*, Vol. 3, No. 2, pp. 31-36.
- [13]. Ningsih, Silvia., Sri Wahyuningsih, & Yuki Novia Nasution., 2016. Perbandingan Kinerja Metode Complete Linkage dan Average Linkage dalam Menentukan Analisis Cluster (Studi Kasus : Produksi Palawiya Provinsi Kalimantan Timur 2014/2015). *Prosiding Seminar Sains dan Teknologi UNMUL*, Vol. 1 thn 2016, 46-50: Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mulawarman, Samarinda.
- [14]. Ramadhani, Lisa., Ika Purnamasari, & Fidia Deny Tisna Amijaya., 2018. Penerapan Metode Complete Linkage dan Metode Hierarchical Clustering Multiscale Bootstrap (Studi Kasus: Kemiskinan di Kalimanta Timur Tahun 2016). *Jurnal EKSPONENSIAL*, Vol. 9, No. 1, pp. 1-10.
- [15]. Rosni., 2017. Analisis Tingkat Kesejahteraan Masyarakat Nelayan di Desa Dahari Selebar Kecamatan Talawi Kabupate Batubara. *Jurnal Geografi*, Vol. 9, No. 1, pp. 53-66.
- [16]. Sapriyanti, & Yan Rianto., 2020. Komparasi Metode Clustering K-Means dan Single Linkage untuk Penentuan Kelompok Agent pada Call Center. *Journal of Information System, Applied, Management, Accounting, and Research*, Vol. 4, No. 3, pp.1-7.
- [17]. Sjafrizal., 2008. *Ekonomi Regional Teori dan Aplikasi*. Baduose Media., Padang.
- [18]. Sukmasari, Dahliana., 2020. Konsep Kesejahteraan Masyarakat dalam Perspektif Al-Qur'an. *Journal of Qur'an and Hadis Studies*, Vo. 3, No. 1, pp. 1-16.
- [19]. Verdian, Edo., 2019. Analisis Faktor yang Merupakan Intensi Perpindahan Merek Transportasi Online di Surabaya. *Jurnal AGORA*, Vol. 7, No. 1, pp. 1-8
- [20]. Widyawati., Wawan Laksito Yuly Saptomo, & Yustina Retno Wahyu Utami., 2020. Penerapan Agglomerative Hierarchical Clustering untuk Segmentasi Pelanggan. *Jurnal Ilmiah Sinus*, Vol. 18, No. 1, pp. 75-87.