

**ANALISIS KLASTER DENGAN *AVERAGE LINKAGE METHOD*
UNTUK PENGELOMPOKAN PROVINSI DI INDONESIA
BERDASARKAN INDIKATOR KESEJAHTERAAN RAKYAT**

***CLUSTERING ANALYSIS WITH AVERAGE LINKAGE METHOD FOR
GROUPING PROVINCE IN INDONESIA BASED ON WELFARE
INDICATORS***

Dwiki Prasetya Subakti¹, Sufri², Gusmi Kholijah³

^{1,2,3} Program Studi Matematika Fakultas Sains dan Teknologi Universitas Jambi

¹dwikiprasetyasubakti@gmail.com, ²sufriary@yahoo.com, ³gusmikholijah@unja.ac.id

ABSTRAK

Tingginya tingkat kesenjangan sosial di Indonesia merupakan masalah yang harus segera diselesaikan. Kesenjangan sosial yang tinggi akan berakibat pada peningkatan tensi sosial yang berdampak pula pada tingginya tingkat konflik dan kriminalitas di masyarakat. Masalah kesenjangan sosial dapat diselesaikan dengan percepatan program pemerataan kesejahteraan oleh pemerintah. Pemberian program ini harus berkeadilan dan disesuaikan dengan kondisi yang diperlukan oleh setiap daerah. Hal ini dikarenakan masing-masing daerah memiliki penyebab masalah kesejahteraan yang berbeda. Oleh karena itu dalam memberikan program tersebut pemerintah harus memiliki skala prioritas terhadap masalah kesejahteraan di suatu wilayah yang dapat dilakukan menggunakan suatu metode matematika di bidang statistika yaitu analisis klaster. Penelitian ini bertujuan untuk memperoleh, menganalisa dan menginterpretasikan hasil pengelompokan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat. Sebanyak 34 Provinsi di Indonesia sebagai objek akan dikelompokkan berdasarkan 20 variabel yang berkaitan dengan kesejahteraan rakyat. Pengelompokan dilakukan dengan menggunakan Metode Hierarki, prosedur pengelompokan Aglomerasi dengan teknik Pautan Rata-rata (*Average Linkage*) dan ukuran Jarak Euklidian. Dari algoritma pengklasteran yang dilakukan diperoleh bahwa dari 34 provinsi di Indonesia terkelompok ke dalam 5 klaster yaitu Klaster 1 terdiri dari 24 anggota yaitu Provinsi Aceh; Sumatera Utara; Sumatera Barat; Riau; Jambi; Sumatera Selatan; Bengkulu; Lampung; Kep.Bangka Belitung; Kep.Riau; Jawa Barat; Jawa Tengah; Jawa Timur; Banten; Nusa Tenggara Barat; Kalimantan Tengah; Kalimantan Selatan; Kalimantan Timur; Kalimantan Utara; Sulawesi Utara; Sulawesi Tengah; Sulawesi Selatan; Sulawesi Tenggara; dan Gorontalo. Klaster 2 terdiri dari 1 anggota yaitu Provinsi DKI Jakarta. Klaster 3 terdiri dari 2 anggota diantaranya yaitu Provinsi DI Yogyakarta dan Bali. Klaster 4 terdiri dari 6 anggota diantaranya yaitu Provinsi Nusa Tenggara Timur; Kalimantan Barat; Sulawesi Barat; Maluku; Maluku Utara; dan Papua Barat. Klaster 5 terdiri dari 1 anggota yaitu Provinsi Papua. Berdasarkan hasil perbandingan nilai rata-rata masing-masing klaster, maka kelima klaster tersebut diurutkan berdasarkan tingkatan kesejahteraannya yaitu: Klaster 3 sebagai klaster sangat baik, klaster 2 sebagai klaster lebih baik, klaster 1 sebagai klaster baik, klaster 4 sebagai klaster cukup baik dan klaster 5 sebagai klaster kurang baik.

Kata Kunci: *Analisis Klaster, Average Linkage, Kesejahteraan Rakyat*

ABSTRACT

The high level of social inequality in Indonesia is a problem that must be resolved immediately. High social inequality will result in an increase in social tension which also impacts on the high level of conflict and crime in society. The problem of social inequality can be solved by accelerating the welfare distribution program by the government. The provision of this program must be fair and adapted to the conditions needed by each region. This is because each region has different causes of welfare problems. Therefore, in providing the program, the government must have a priority scale on welfare issues in an area that can be done using a mathematical method in the field of statistics, namely cluster analysis. This study aims to obtain, analyze and interpret the results of grouping provinces in Indonesia based on indicators of people's welfare. As many as 34 provinces in Indonesia as objects will be grouped based on 20 variables related to people's welfare. The grouping is done using the Hierarchy Method, the agglomeration grouping procedure with the Average Linkage technique and the size of the Euclidean Distance. From the clustering algorithm, it was found that from 34 provinces in Indonesia grouped into 5 clusters namely Cluster 1 consisting of 24 members namely Aceh Province; North Sumatra; West Sumatra; Riau; Jambi; South Sumatra; Bengkulu; Lampung; Head of Pacific Islands; Riau Islands; West Java; Central Java; East Java; Banten; West Nusa Tenggara; Central Kalimantan; South Borneo; East Kalimantan; North Kalimantan; North Sulawesi; Central Sulawesi; South Sulawesi; Southeast Sulawesi; and Gorontalo. Cluster 2 consists of 1 member, DKI Jakarta Province. Cluster 3 consists of 2 members including DI Yogyakarta and Bali Provinces. Cluster 4 consists of 6 members including East Nusa Tenggara Province; West Kalimantan; West Sulawesi; Maluku; North Maluku; and West Papua. Cluster 5 consists of 1 member, Papua Province. Based on the comparison of the average value of each cluster, the five clusters are sorted based on their level of welfare, namely: Cluster 3 as a very good cluster, Cluster 2 as a better cluster, Cluster 1 as a good cluster, Cluster 4 as a pretty good cluster and cluster 5 as a less good cluster.

Keywords: Cluster Analysis, Average Linkage, People's Welfare

I. PENDAHULUAN

Kesenjangan sosial masih menjadi masalah pelik di Indonesia. Berbagai kebijakan telah diupayakan oleh pemerintah, akan tetapi masih belum terlaksana sesuai dengan yang diharapkan. Salah satu penyebab kegagalan kebijakan tersebut adalah kurangnya pendalaman pemerintah dalam meninjau masalah kemiskinan dan kesejahteraan di suatu daerah. Kebijakan pemerintah menjadi tidak efisien dan tidak tepat sasaran dikarenakan setiap daerah memiliki penyebab masalah kemiskinan dan kesejahteraan yang berbeda. Oleh karena itu, untuk mengatasi kesenjangan sosial dan mewujudkan kesejahteraan yang merata pemerintah harus punya skala prioritas atas kondisi masyarakat yang sangat memerlukan peningkatan dalam hal kesejahteraan sosial.

Upaya yang dapat dilakukan dalam mengenali keadaan dan masalah suatu masyarakat di daerah tertentu agar kebijakan yang diberikan dapat tepat sasaran adalah dengan melakukan pengklasifikasian wilayah berdasarkan kondisi tingkat kesejahteraan masyarakat. Pengelompokan wilayah tersebut dapat dilakukan menggunakan suatu metode matematika di bidang statistika yaitu analisis kluster (*clustering analysis*)[5]. Pada prinsipnya analisis kluster merupakan metode yang digunakan untuk pengelompokan objek, dimana objek dikelompokkan ke dalam sebuah kluster atau kelompok yang mempunyai ciri relatif sama, sedangkan antar kluster atau kelompok memiliki ciri yang berbeda [7].

Permasalahan yang akan dikaji dalam penelitian ini adalah bagaimana pengelompokan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat dengan menggunakan analisis kluster *average linkage*. Penelitian ini bertujuan untuk memperoleh, menganalisa dan menginterpretasikan hasil pengelompokan provinsi di Indonesia berdasarkan indikator kesejahteraan rakyat. Penelitian ini ditujukan kepada pemerintah pusat dan daerah sebagai sumber informasi tentang pengelompokan provinsi Indonesia berdasarkan indikator kesejahteraan rakyat. Sehingga pemerintah dapat melakukan kebijakannya secara bijak dan tepat sasaran, yang berdampak kepada penurunan tingkat kesenjangan dan pemerataan kesejahteraan rakyat di seluruh Indonesia. Batasan masalah penelitian ini yaitu: (1) Objek yang akan dikelompokkan adalah seluruh Provinsi di Indonesia sebanyak 34 Provinsi. (2) Variabel yang digunakan sebanyak 20 variabel yang berkaitan dengan indikator kesejahteraan rakyat di Indonesia tahun 2017, diperoleh dari Badan Pusat Statistik. (3) Metode analisis kluster yang digunakan adalah Metode Hierarki (*Hierarchical Method*) menggunakan prosedur pengelompokan Aglomerasi (*Agglomerative*) dengan teknik Pautan Rata-rata (*Average Linkage*) dan ukuran Jarak Euklidean (*Euclidean Distance*). (4) Proses pengelompokan pada penelitian ini dikerjakan secara manual dan hasilnya dibandingkan dengan output pengerjaan pada *software R*. (5) Hasil pengelompokan pada penelitian ini diurutkan dan diberi nama berdasarkan tingkat kesejahteraan dengan melakukan perbandingan nilai rata-rata antar kluster.

II. METODE

Jenis data yang digunakan pada penelitian ini adalah data sekunder yang berkaitan dengan Kesejahteraan Rakyat pada Tahun 2017. Data dikumpulkan dan diperoleh melalui website Badan Pusat Statistik (BPS) Indonesia (<https://bps.go.id>), Publikasi BPS “Indikator

Kesejahteraan Rakyat”[2], Publikasi BPS “Statistik Kesejahteraan Rakyat”[4], dan Publikasi BPS “Statistik Indonesia”[3].

Objek yang dikelompokkan pada penelitian ini adalah Provinsi di Indonesia sebanyak 34 Provinsi. Variabel yang digunakan adalah 20 variabel diantaranya: (V1) Kepadatan Penduduk mewakili kesejahteraan di bidang kependudukan (*population*); (V2) Angka Harapan Hidup, (V3) Angka Kesakitan mewakili kesejahteraan di bidang kesehatan dan gizi (*health and nutrition*); (V4) Persentase Penduduk yang Melek Huruf, (V5) Angka Partisipasi Murni SD/MI, (V6) Angka Partisipasi Murni SMP/MTs, (V7) Angka Partisipasi Murni SMA/SMK/MA mewakili kesejahteraan di bidang pendidikan (*education*); (V8) Tingkat Pengangguran Terbuka (V9) Tingkat Partisipasi Angkatan Kerja mewakili kesejahteraan di bidang Ketenagakerjaan (*employment*); (V10) Rata-rata Pengeluaran Untuk Bahan Non-Makanan, (V11) Rata-rata Konsumsi Kalori Perhari, (V12) Rata-rata Konsumsi Protein Perhari mewakili kesejahteraan di bidang taraf dan pola konsumsi (*consumption level and pattern*); (V13) Persentase Rumah dengan Status Milik Sendiri, (V14) Persentase Rumah dengan Penerangan Listrik PLN, (V15) Persentase Rumah dengan Sumber Air Minum Layak, (V16) Persentase Rumah dengan Sanitasi Layak mewakili kesejahteraan di bidang perumahan dan lingkungan (*housing and environment*); (V17) Persentase Penduduk Miskin mewakili kesejahteraan di bidang kemiskinan (*poverty*); (V18) Persentase Penduduk yang melakukan Bepergian Wisata, (V19) Persentase Penduduk yang Menjadi Korban Kejahatan dan (V20) Persentase Penduduk yang Menggunakan Telepon Seluler mewakili kesejahteraan di bidang sosial lainnya (*other social concern*).

Metode analisis kluster yang digunakan pada penelitian ini adalah Metode Hierarki (*Hierarchical Method*) menggunakan prosedur pengelompokan Aglomerasi (*Agglomerative*) dengan Teknik Pautan Rata-rata (*Average Linkage*) dan ukuran Jarak Euklidean (*Euclidean Distance*). Langkah-langkah pengelompokan pada penelitian ini adalah sebagai berikut:

1. Merumuskan masalah

Menentukan objek yang akan dikelompokkan dan variabel yang digunakan.

2. Melakukan standarisasi data

Standarisasi data dilakukan agar variabel yang digunakan mempunyai satuan yang sama. Proses standarisasi adalah dengan melakukan transformasi data awal dari masing-masing variabel ke dalam bentuk *Z-Score*. Transformasi data ke dalam bentuk *Z-Score* menghasilkan data pengamatan baru yang memenuhi distribusi peubah acak normal dengan rata-rata nol dan

variansi satu atau dapat disebut sebagai distribusi Normal Baku. Z-skor dapat dihitung dengan rumus[1]:

$$Z_{ij} = \frac{x_{ij} - \mu}{\sigma} \tag{1}$$

dengan:

Z_{ij} = Z-Score objek ke- i observasi ke- j

x_{ij} = nilai objek ke- i dan observasi ke- j

μ = rata-rata observasi

σ = simpangan baku observasi dengan

3. Menentukan ukuran kesamaan antar objek

Ukuran jarak ditentukan menggunakan Jarak Euklidean dengan persamaan berikut[6]:

$$d(i, j) = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2} \tag{2}$$

dengan:

$d(i, j)$ = Jarak Euklidean antara objek ke- i dan objek ke- j

X_{ik} = nilai dari objek ke- i pada variabel ke- k

X_{jk} = nilai dari objek ke- j pada variabel ke- k

p = banyaknya variabel yang diamati

Keseluruhan objek secara berpasangan dihitung jarak kesamaannya untuk memperoleh matriks jarak awal yang digunakan pada tahap pengklasteran. Matriks Jarak dinyatakan dengan:

$$d(i, j) = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1p} \\ d_{21} & d_{22} & \dots & d_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & d_{np} \end{pmatrix}$$

Matriks Jarak Euklidean $d(i, j)$ menyatakan jarak antara objek ke- i dan objek ke- j . Matriks Jarak Euklidean tersebut mempunyai sifat-sifat diantaranya yaitu[6]:

- a. $d(i, i) = 0$
- b. Simetris, $d(i, j) = d(j, i)$
- c. Positif, $d(i, j) \geq 0$
- d. $d(i, j)$ akan meningkat nilainya seiring tidak miripnya antara objek i dan j .

Jarak antar objek tersebut memiliki arti bahwa semakin kecil nilai jarak antar objek (d) maka semakin dekat atau mirip kedua objek tersebut dan semakin besar nilai jarak antar objek maka semakin jauh atau tidak mirip kedua objek tersebut.

4. Proses pengklasteran dengan metode pautan rata-rata (*average linkage*).

Algoritma pengklasteran penelitian ini adalah sebagai berikut [6]:

- a. Menggabungkan dua objek dengan jarak minimum dalam $D = \{d_{ij}\}$
- b. Menghitung jarak antara kluster yang telah terbentuk pada langkah (a) dengan provinsi lainnya
- c. Dari algoritma di atas jarak-jarak antara (ij) dan kluster k yang lain dihitung dengan:

$$d_{C_i, C_k} = \text{average}(d_{ik}, d_{jk})$$

Tahapan dilanjutkan dengan membentuk matriks jarak baru dan mengulangi langkah ke (a) dan seterusnya hingga membentuk suatu tingkatan (hierarki) kluster dan tergambarkan melalui dendogram.

5. Menentukan jumlah kluster dan anggota kluster berdasarkan hasil pengolahan data.
6. Melakukan interpretasi terhadap hasil pengelompokkan/kluster yang diperoleh.

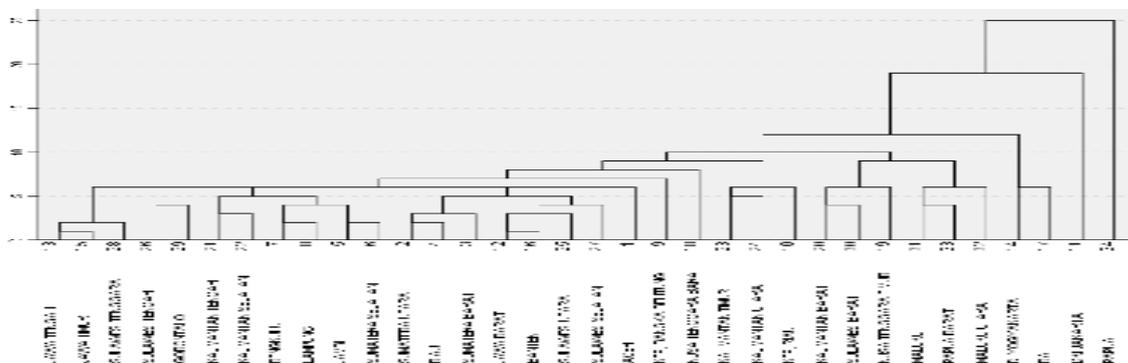
III. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini memiliki perbedaan satuan yang sangat bervariasi sehingga proses pengklasteran tidak dapat langsung dilakukan. Pada penelitian ini proses standarisasi data dilakukan dengan melakukan transformasi pada data asli menjadi *Z-Score* sebelum dilanjutkan ke tahapan selanjutnya. Standarisasi data dilakukan untuk keseluruhan variabel yang digunakan dengan menggunakan persamaan (1).

Setelah data distandarisasi, selanjutnya adalah menghitung ukuran kesamaan antar objek dengan Jarak Euklidean pada persamaan (2) dan membentuk matriks jarak awal. Proses perhitungan jarak dilakukan hingga seluruh pasangan dua objek telah diketahui jaraknya. Perhitungan ini akan menghasilkan matriks jarak awal yang akan digunakan untuk melakukan algoritma pengklasteran.

Proses pengklasteran dimulai dari jarak minimum/terdekat antar objek pada matriks jarak. Selanjutnya dihitung jarak antara kluster baru yang terbentuk dengan objek lainnya

dengan *average linkage* dan melakukan perubahan pada matriks jarak untuk dilanjutkan ke iterasi berikutnya. Diperoleh hasil pengklasteran dalam *dendogram* pada Gambar 1 berikut:



Gambar 1. Dendogram Hasil Pengklasteran Provinsi di Indonesia

Teknis pengelompokan tersebut dimulai dari dua objek yang memiliki jarak terdekat, dapat dijelaskan dengan *Amalgamation Steps* berikut:

1. Tahap 1 objek 13 dan 15 bergabung menjadi kelompok 13 dengan jarak minimum sebesar 1,641.
2. Tahap 2 objek 12 dan 16 bergabung menjadi kelompok 12 dengan jarak minimum sebesar 1,787.
3. Tahap 3 objek 2 dan 4 bergabung menjadi kelompok 2 dengan jarak minimum sebesar 2,121.
4. Tahap 4 objek 7 dan 8 bergabung menjadi kelompok 7 dengan jarak minimum sebesar 2,402.
5. Tahap 5 objek 5 dan 6 bergabung menjadi kelompok 5 dengan jarak minimum sebesar 2,409.
6. Tahap 6 objek 28 bergabung dengan kelompok 13 dengan jarak minimum sebesar 2,508 menjadi kelompok 13 yang terdiri dari 3 objek (13,15,dan 28).
7. Tahap 7 objek 3 bergabung dengan kelompok 2 dengan jarak minimum sebesar 2,642 menjadi kelompok 2 yang terdiri dari 3 objek (2,3 dan 4).
8. Tahap 8 objek 25 bergabung dengan kelompok 12 dengan jarak minimum sebesar 2,687 menjadi kelompok 12 yang terdiri dari 3 objek (12,16 dan 25).
9. Tahap 9 objek 21 dan 22 bergabung menjadi kelompok 21 dengan jarak minimum sebesar 2,772.

10. Tahap 10 kelompok 7 bergabung dengan kelompok 5 dengan jarak minimum sebesar 3,080 menjadi kelompok 5 yang terdiri dari 4 objek (5,6,7 dan 8).
11. Tahap 11 objek 31 dan 33 bergabung menjadi kelompok 31 dengan jarak minimum sebesar 3,124.
12. Tahap 12 objek 27 bergabung dengan kelompok 12 dengan jarak minimum sebesar 3,137 menjadi kelompok 12 yang terdiri dari 4 objek (12,16, 25, 27).
13. Tahap 13 objek 20 dan 30 bergabung menjadi kelompok 20 dengan jarak minimum sebesar 3,156.
14. Tahap 14 objek 26 dan 29 bergabung menjadi kelompok 26 dengan jarak minimum sebesar 3,196.
15. Tahap 15 objek 23 dan 24 bergabung menjadi kelompok 23 dengan jarak minimum sebesar 3,515.
16. Tahap 16 kelompok 2 bergabung dengan kelompok 12 dengan jarak minimum sebesar 3,610 menjadi kelompok 2 yang terdiri dari 7 objek (2,3,4,12,16,25 dan 27).
17. Tahap 17 kelompok 5 bergabung dengan kelompok 21 dengan jarak minimum sebesar 3,736 menjadi kelompok 5 yang terdiri dari 6 objek (5,6,7,8,21 dan 22).
18. Tahap 18 objek 14 dan 17 bergabung menjadi kelompok 14 dengan jarak minimum sebesar 3,879.
19. Tahap 19 kelompok 13 bergabung dengan kelompok 26 dengan jarak minimum sebesar 3,884 menjadi kelompok 13 yang terdiri dari 5 objek (13,15,26,28 dan 29).
20. Tahap 20 objek 10 bergabung dengan kelompok 23 dengan jarak minimum sebesar 4,015 menjadi kelompok 10 yang terdiri dari 3 objek (10, 23, 24).
21. Tahap 21 objek 1 bergabung dengan kelompok 2 dengan jarak minimum sebesar 4,068 menjadi kelompok 1 yang terdiri dari 8 objek (1,2,3,4,12,16,25 dan 27).
22. Tahap 22 objek 19 bergabung dengan kelompok 20 dengan jarak minimum sebesar 4,121 menjadi kelompok 19 yang terdiri dari 3 objek (19,20 dan 30).
23. Tahap 23 kelompok 13 bergabung dengan kelompok 5 dengan jarak minimum sebesar 4,197 menjadi kelompok 5 yang terdiri dari 11 objek (5,6,7,8,13,15,21,22,26,28, dan 29).
24. Tahap 24 objek 32 bergabung dengan kelompok 31 dengan jarak minimum sebesar 4,198 menjadi kelompok 31 yang terdiri dari 3 objek (31,32 dan 33).

25. Tahap 25 kelompok 1 bergabung dengan kelompok 5 dengan jarak minimum sebesar 4,301 menjadi kelompok 1 yang terdiri dari 19 objek (1,2,3,4,5,6,7,8,12,13,15,16,21,22,25,26,27,28,dan 29).
26. Tahap 26 objek 9 bergabung dengan kelompok 1 dengan jarak minimum sebesar 4,435 menjadi kelompok 1 yang terdiri dari 20 objek (1,2,3,4,5,6,7,8,9,12,13,15,16,21,22,25,26,27,28,dan 29).
27. Tahap 27 objek 18 bergabung dengan kelompok 1 dengan jarak minimum sebesar 5,186 menjadi kelompok 1 yang terdiri dari 21 objek (1,2,3,4,5,6,7,8,9,12,13,15,16,18,21,22,25,26,27,28,dan 29).
28. Tahap 28 kelompok 1 bergabung dengan kelompok 10 dengan jarak minimum sebesar 5,318 menjadi kelompok 1 yang terdiri dari 24 objek (1,2,3,4,5,6,7,8,9,10,12,13,15,16,18,21,22,23,24,25,26,27,28,dan 29).
29. Tahap 29 kelompok 19 bergabung dengan kelompok 31 dengan jarak minimum sebesar 5,394 menjadi kelompok 19 yang terdiri dari 6 objek (19,20,30,31,32, dan 33).
30. Tahap 30 kelompok 1 bergabung dengan kelompok 19 dengan jarak minimum sebesar 5,865 menjadi kelompok 1 yang terdiri dari 30 objek (1,2,3,4,5,6,7,8,9,10,12,13,15,16,18,19,20,21,22,23,24,25,26,27,28, 29,30,31,32 dan 33).
31. Tahap 31 kelompok 1 bergabung dengan kelompok 14 dengan jarak minimum sebesar 6,640 menjadi kelompok 1 yang terdiri dari 32 objek (1,2,3,4,5,6,7,8,9,10,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26, 27,28,29,30,31,32 dan 33).
32. Tahap 32 objek 11 bergabung dengan kelompok 1 dengan jarak minimum sebesar 9,780 menjadi kelompok 1 yang terdiri dari 33 objek (1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28, 29,30,31,32 dan 33).
33. Tahap 33 objek 34 bergabung dengan kelompok 1 dengan jarak minimum sebesar 12,748 menjadi kelompok 1 yang terdiri dari 34 objek (1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28, 29,30,31,32,33,34).

Berdasarkan hasil pengklasteran tersebut disimpulkan bahwa dari 34 objek penelitian yaitu provinsi di Indonesia dibentuk ke dalam 5 klaster dengan anggota untuk masing-masing klaster disajikan dalam Tabel 1 berikut:

Tabel 1. Anggota Klaster yang Terbentuk dengan *Average Linkage*

Klaster	Anggota Klaster	Jumlah
1	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kep. Riau, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, dan Gorontalo.	24 anggota
2	DKI Jakarta	1 anggota
3	DI Yogyakarta dan Bali	2 anggota
4	Nusa Tenggara Timur, Kalimantan Barat, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat	6 anggota
5	Papua	1 anggota

Selanjutnya, untuk mengetahui karakteristik setiap klaster digunakan perhitungan dengan membandingkan nilai rata-rata untuk masing-masing variabel dalam masing-masing klasternya[1]. Data rata-rata hasil klaster yang terbentuk dapat dilihat pada Tabel 2 sebagai berikut:

Tabel 2. Perhitungan Rata-rata Variabel Tiap Klaster

Variabel	Klaster 1	Klaster 2	Klaster 3	Klaster 4	Klaster 5
V1 per km ²	285,38	15624,00	968,00	50,83	10,00
V2 tahun	69,90	72,55	73,10	66,43	65,14
V3 (%)	13,93	12,34	15,60	13,18	8,43
V4 (%)	97,14	99,68	94,34	95,84	75,80
V5 (%)	97,33	97,64	97,69	95,60	78,83
V6 (%)	77,86	80,72	84,29	70,15	56,13
V7 (%)	62,10	59,54	71,03	58,61	43,48
V8 (%)	5,14	6,25	2,16	5,21	3,79
V9 (%)	68,25	66,08	73,91	68,11	77,12
V10 rupiah	503898,36	1199617,99	706417,31	408388,09	441507,44
V11 kkal	2155,90	2153,62	2246,77	1932,49	1924,39
V12 gram	62,10	68,44	66,95	54,06	46,03
V13 (%)	78,58	48,33	72,90	83,75	81,00
V14 (%)	94,64	99,80	99,85	79,20	41,61
V15 (%)	69,99	88,93	84,02	66,97	59,09
V16 (%)	65,93	91,13	89,95	58,20	33,06
V17 (%)	10,02	3,78	8,44	14,93	27,69
V18 (%)	26,52	43,04	41,35	18,35	12,26
V19 (%)	1,29	1,20	0,99	1,23	1,89
V20 (%)	74,81	82,77	75,14	65,35	42,59

Berdasarkan perbandingan nilai rata-rata variabel untuk setiap klaster maka dapat diketahui karakteristik dari kelima klaster yang terbentuk sebagai berikut:

1. Klaster 1

Berdasarkan perbandingan nilai rata-rata antar klaster maka diperoleh karakteristik klaster 1 yang tersaji dalam Tabel 3 berikut:

Tabel 3. Karakteristik Klaster 1

Kategori	Variabel-variabel yang memenuhi		
Sangat Baik	-		
Lebih Baik	V4	Persentase Penduduk yang Melek Huruf	97,14 %
	V7	Angka Partisipasi Murni SMA/SMK/MA	62,10 %
	V11	Rata-rata Konsumsi Kalori perhari	2155,90 kkal
Baik	V1	Kepadatan Penduduk per Km2	285,38
	V2	Angka Harapan Hidup	penduduk/km ²
	V5	Angka Partisipasi Murni SD/MI	69,90 tahun
	V6	Angka Partisipasi Murni SMP/MTs	97,33 %
	V8	Tingkat Pengangguran Terbuka	77,86 %
	V9	Tingkat Partisipasi Angkatan Kerja	5,14 %
	V10	Rata-rata Pengeluaran Untuk Bahan Non-Makanan	68,25 %
	V12	Rata-rata Konsumsi Protein perhari	Rp. 503.898,36
	V13	Persentase Rumah dengan Status Milik Sendiri	62,0 gram
	V14	Persentase Rumah dengan Penerangan Listrik PLN	78,58 %
	V15	Persentase Rumah dengan Sumber Air Minum Layak	94,64 %
	V16	Persentase Rumah dengan Sanitasi Layak	69,99 %
	V17	Persentase Penduduk Miskin	65,93 %
	V18	Persentase Penduduk yang Melakukan Bepergian	10,02 %
Cukup Baik	V3	Angka Kesakitan	13,93 %
	V19	Persentase Penduduk Menjadi Korban Kejahatan	1,29 %
Kurang Baik	-		

2. Klaster 2

Berdasarkan perbandingan nilai rata-rata antar klaster maka diperoleh karakteristik klaster 2 yang tersaji dalam Tabel 4 berikut:

Tabel 4. Karakteristik Klaster 2

Kategori	Variabel-variabel yang memenuhi		
Sangat Baik	V4	Persentase Penduduk yang Melek Huruf	99,68 %
	V10	Rata-rata Pengeluaran Untuk Bahan Non-Makanan	Rp. 1.199.617,99
	V12	Rata-rata Konsumsi Protein perhari	68,44 gram
	V15	Persentase Rumah dengan Sumber Air Minum Layak	88,93 %
	V16	Persentase Rumah dengan Sanitasi Layak	91,13 %
	V17	Persentase Penduduk Miskin	3,78 %
	V18	Persentase Penduduk yang Melakukan Bepergian	43,04 %
	V18	Wisata	82,77 %
	V20	Persentase Penduduk yang Menggunakan Ponsel	
Lebih Baik	V2	Angka Harapan Hidup	72,55 tahun
	V3	Angka Kesakitan	12,34 %
	V5	Angka Partisipasi Murni SD/MI	97,64 %
	V6	Angka Partisipasi Murni SMP/MTs	80,72 %
	V14	Persentase Rumah dengan Penerangan Listrik PLN	99,80 %
Baik	V19	Persentase Penduduk Menjadi Korban Kejahatan	1,20 %
	V7	Angka Partisipasi Murni SMA/SMK/MA	59,54 %
Cukup Baik	V11	Rata-rata Konsumsi Kalori perhari	2153,62 kkal
	-		

Kurang Baik	V1	Kepadatan Penduduk per Km ²	15.624 penduduk/km ²
	V8	Tingkat Pengangguran Terbuka	6,25 %
	V9	Tingkat Partisipasi Angkatan Kerja	66,08 %
	V13	Persentase Rumah dengan Status Milik Sendiri	48,33 %

3. Kluster 3

Berdasarkan perbandingan nilai rata-rata antar kluster maka diperoleh karakteristik kluster 3 yang tersaji dalam Tabel 5 berikut:

Tabel 5. Karakteristik Kluster 3

Kategori	Variabel-variabel yang memenuhi		
Sangat Baik	V2	Angka Harapan Hidup	73,10 tahun
	V5	Angka Partisipasi Murni SD/MI	97,69 %
	V6	Angka Partisipasi Murni SMP/MTs	84,29 %
	V7	Angka Partisipasi Murni SMA/SMK/MA	71,03 %
	V8	Tingkat Pengangguran Terbuka	2,16 %
	V11	Rata-rata Konsumsi Kalori perhari	2246,77 kkal
	V14	Persentase Rumah dengan Penerangan Listrik PLN	99,85 %
	V19	Persentase Penduduk Menjadi Korban Kejahatan	0,99 %
Lebih Baik	V9	Tingkat Partisipasi Angkatan Kerja	73,91 %
	V10	Rata-rata Pengeluaran Untuk Bahan Non-Makanan	Rp. 706.417,31
	V12	Rata-rata Konsumsi Protein perhari	66,95 gram
	V15	Persentase Rumah dengan Sumber Air Minum Layak	84,02 %
	V16	Persentase Rumah dengan Sanitasi Layak	89,95 %
	V17	Persentase Penduduk Miskin	8,44 %
	V18	Persentase Penduduk yang Melakukan Bepergian Wisata	41,35 %
	V20	Persentase Penduduk yang Menggunakan Ponsel	75,14 %
Baik	-		
Cukup Baik	V1	Kepadatan Penduduk per Km ²	968 penduduk/km ²
	V4	Persentase Penduduk yang Melek Huruf	94,34 %
Kurang Baik	V13	Persentase Rumah dengan Status Milik Sendiri	72,90 %
	V3	Angka Kesakitan	15,60 %

4. Kluster 4

Berdasarkan perbandingan nilai rata-rata antar kluster maka diperoleh karakteristik kluster 4 yang tersaji dalam Tabel 6 berikut:

Tabel 6. Karakteristik Kluster 4

Kategori	Variabel-variabel yang memenuhi		
Sangat Baik	V13	Persentase Rumah dengan Status Milik Sendiri	83,75%
Lebih Baik	V1	Kepadatan Penduduk per Km ²	50,83 penduduk/km ²
	V3	Angka Kesakitan	13,18 %
Baik	V4	Persentase Penduduk yang Melek Huruf	95,84 %
	V19	Persentase Penduduk Menjadi Korban Kejahatan	1,23 %
Cukup Baik	V2	Angka Harapan Hidup	66,43 tahun
	V5	Angka Partisipasi Murni SD/MI	95,60 %
	V6	Angka Partisipasi Murni SMP/MTs	70,15 %
	V7	Angka Partisipasi Murni SMA/SMK/MA	58,61 %
	V8	Tingkat Pengangguran Terbuka	5,21 %

	V9	Tingkat Partisipasi Angkatan Kerja	68,11 %
	V11	Rata-rata Konsumsi Kalori perhari	1932,49 kkal
	V12	Rata-rata Konsumsi Protein perhari	54,05 gram
	V14	Persentase Rumah dengan Penerangan Listrik PLN	79,20 %
	V15	Persentase Rumah dengan Sumber Air Minum Layak	66,97 %
	V16	Persentase Rumah dengan Sanitasi Layak	58,20 %
	V17	Persentase Penduduk Miskin	14,93 %
	V18	Persentase Penduduk yang Melakukan Bepergian Wisata	18,35 %
	V20	Persentase Penduduk yang Menggunakan Ponsel	65,35 %
Kurang Baik	V10	Rata-rata Pengeluaran Untuk Bahan Non-Makanan	Rp. 408.388,09

5. Klaster 5

Berdasarkan perbandingan nilai rata-rata antar klaster maka diperoleh karakteristik klaster 5 yang tersaji dalam Tabel 7 berikut:

Tabel 7. Karakteristik Klaster 5

Kategori	Variabel-variabel yang memenuhi		
Sangat Baik	V1	Kepadatan Penduduk per Km ²	10 penduduk/km ²
	V3	Angka Kesakitan	8,43 %
	V9	Tingkat Partisipasi Angkatan Kerja	77,12 %
Lebih Baik	V8	Tingkat Pengangguran Terbuka	3,79 %
	V13	Persentase Rumah dengan Status Milik Sendiri	81,00 %
Baik	-		
Cukup Baik	V10	Rata-rata Pengeluaran Untuk Bahan Non-Makanan	Rp. 441.507,44
Kurang Baik	V2	Angka Harapan Hidup	65,14 tahun
	V4	Persentase Penduduk yang Melek Huruf	75,80 %
	V5	Angka Partisipasi Murni SD/MI	78,83 %
	V6	Angka Partisipasi Murni SMP/MTs	56,13 %
	V7	Angka Partisipasi Murni SMA/SMK/MA	43,48 %
	V11	Rata-rata Konsumsi Kalori perhari	1924,39 kkal
	V12	Rata-rata Konsumsi Protein perhari	46,03 gram
	V14	Persentase Rumah dengan Penerangan Listrik	41,61 %
	V15	Persentase Rumah dengan Sumber Air Minum Layak	59,09 %
	V16	Persentase Rumah dengan Sanitasi Layak	33,06 %
	V17	Persentase Penduduk Miskin	27,69 %
	V18	Persentase Penduduk yang Melakukan Bepergian Wisata	12,26 %
	V19	Persentase Penduduk Menjadi Korban Kejahatan	1,89 %
V20	Persentase Penduduk yang Menggunakan Ponsel	42,95 %	

Berdasarkan karakteristik dari masing-masing klaster yang telah dijelaskan di atas serta melalui analisis dan pertimbangan pada 20 variabel yang digunakan, maka klaster-klaster yang telah terbentuk diurutkan dan diberi nama yang mewakili klaster tersebut.

Tabel 8. Tingkatan Klaster yang Terbentuk

Kategori	Anggota Klaster	Klaster
Sangat Baik	DI Yogyakarta dan Bali	3
Lebih Baik	DKI Jakarta	2
Baik	Aceh, Sumatera Utara, Sumatera Barat, Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep.Bangka Belitung, Kep.Riau, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Kalimantan	1

	Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, dan Gorontalo.	
Cukup Baik	Nusa Tenggara Timur, Kalimantan Barat, Sulawesi Barat, Maluku, Maluku Utara, Papua Barat	4
Kurang Baik	Papua	5

Klaster provinsi yang telah terbentuk dapat digambarkan pada peta Indonesia yang tersaji dalam Gambar 2 berikut :



Gambar 2. Peta Hasil Pengklasteran Provinsi di Indonesia

Berdasarkan analisis di atas dapat dijelaskan bahwa: Klaster pertama beranggotakan 24 provinsi (lihat Tabel 1). Klaster 1 memiliki karakteristik yaitu tidak satupun variabel memenuhi kategori sangat baik, 3 variabel dengan kategori lebih baik, 15 variabel dengan kategori baik serta 2 variabel dengan kategori cukup baik jika dibandingkan dengan keempat klaster lainnya (lihat Tabel 3). Jika dianalisis dari masing-masing variabel kesejahteraan yang digunakan, klaster pertama ini tergolong ke dalam klaster yang baik. Hal ini dikarenakan dari kedua puluh variabel yang digunakan, cenderung tergolong ke dalam kategori baik.

Klaster kedua beranggotakan 1 provinsi yaitu Provinsi DKI Jakarta (lihat Tabel 1). Klaster 2 memiliki karakteristik yaitu 8 variabel dengan kategori sangat baik, 6 variabel dengan kategori lebih baik, 2 variabel dengan kategori baik serta 4 variabel dengan kategori kurang baik jika dibandingkan dengan keempat klaster lainnya (lihat Tabel 4). Jika dianalisis dari masing-masing variabel kesejahteraan yang digunakan, klaster kedua ini tergolong ke dalam klaster yang berada pada tingkat lebih baik. Hal ini dikarenakan dari kedua puluh

variabel yang digunakan, cenderung tergolong ke dalam kategori sangat baik, lebih baik dan baik. Akan tetapi klaster 2 masih terdapat 4 variabel dalam kategori kurang baik, sehingga menyebabkan klaster tersebut belum bisa mencapai sangat baik.

Klaster ketiga beranggotakan 2 provinsi yaitu Provinsi DI Yogyakarta dan Provinsi Bali (lihat Tabel 1). Klaster 3 memiliki karakteristik yaitu 8 variabel dengan kategori sangat baik, 8 variabel dengan kategori lebih baik, 3 variabel dengan kategori cukup baik serta 1 variabel dengan kategori kurang baik jika dibandingkan dengan keempat klaster lainnya (lihat Tabel 5). Jika dianalisis dari masing-masing variabel kesejahteraan yang digunakan, klaster ketiga ini tergolong ke dalam klaster yang sangat baik. Hal ini dikarenakan dari kedua puluh variabel yang digunakan, cenderung tergolong ke dalam kategori sangat baik, lebih baik dan cukup baik. Klaster 3 tersebut juga tidak memiliki banyak variabel dalam kategori kurang baik, sehingga menyebabkan klaster tersebut bisa dikatakan sangat baik.

Klaster keempat beranggotakan 6 provinsi (lihat Tabel 1). Klaster 4 memiliki karakteristik yaitu 1 variabel memenuhi kategori sangat baik, 1 variabel dengan kategori lebih baik, 3 variabel dengan kategori baik, 14 variabel dengan kategori cukup baik serta 1 variabel dengan kategori kurang baik jika dibandingkan dengan keempat klaster lainnya (lihat Tabel 6). Jika dianalisis dari masing-masing variabel kesejahteraan yang digunakan, klaster keempat ini tergolong ke dalam klaster yang cukup baik. Hal ini dikarenakan dari kedua puluh variabel yang digunakan cenderung tergolong ke dalam kategori cukup baik.

Klaster kelima terdiri dari 1 provinsi yaitu Papua (lihat Tabel 1). Klaster 5 memiliki karakteristik yaitu 3 variabel memenuhi kategori sangat baik, 2 variabel dengan kategori lebih baik, 1 variabel dengan kategori cukup baik serta 14 variabel dengan kategori kurang baik jika dibandingkan dengan keempat klaster lainnya (lihat Tabel 7). Jika dianalisis dari masing-masing variabel kesejahteraan yang digunakan, klaster keempat ini tergolong ke dalam klaster yang kurang baik. Hal ini dikarenakan dari kedua puluh variabel yang digunakan cenderung tergolong ke dalam kategori yang kurang baik atau berada di bawah keempat klaster lainnya.

IV. KESIMPULAN DAN SARAN

Terdapat 5 klaster dari 34 Provinsi yang dikelompokkan. Adapun anggota dari masing-masing klaster yang terbentuk adalah sebagai berikut:

1. Klaster 1 terdiri dari 24 anggota diantaranya yaitu Provinsi (1) Aceh; (2) Sumatera Utara; (3) Sumatera Barat; (4) Riau; (5) Jambi; (6) Sumatera Selatan; (7) Bengkulu; (8) Lampung; (9) Kep.Bangka Belitung; (10) Kep.Riau; (11) Jawa Barat; (12) Jawa Tengah; (13) Jawa Timur; (14) Banten; (15) Nusa Tenggara Barat; (16) Kalimantan Tengah; (17) Kalimantan Selatan; (18) Kalimantan Timur; (19) Kalimantan Utara; (20) Sulawesi Utara; (21) Sulawesi Tengah; (22) Sulawesi Selatan; (23) Sulawesi Tenggara; dan (24) Gorontalo. Klaster 1 dinamakan sebagai klaster dengan tingkat kesejahteraan Baik.

2. Klaster 2 terdiri dari 1 anggota yaitu Provinsi DKI Jakarta. Klaster 2 dinamakan sebagai klaster dengan tingkat kesejahteraan Lebih Baik.
3. Klaster 3 terdiri dari 2 anggota diantaranya yaitu Provinsi DI Yogyakarta dan Bali. Klaster 3 dinamakan sebagai klaster dengan tingkat kesejahteraan Sangat Baik.
4. Klaster 4 terdiri dari 6 anggota diantaranya yaitu Provinsi (1) Nusa Tenggara Timur; (2) Kalimantan Barat; (3) Sulawesi Barat; (4) Maluku; (5) Maluku Utara; dan (6) Papua Barat. Klaster 4 dinamakan sebagai klaster dengan tingkat kesejahteraan Cukup Baik.
5. Klaster 5 terdiri dari 1 anggota yaitu Provinsi Papua. Klaster 5 dinamakan sebagai klaster dengan tingkat kesejahteraan Kurang Baik.

V. UCAPAN TERIMA KASIH

Terima kasih kepada Kementrian Riset Teknologi dan Pendidikan Tinggi atas beasiswa Bidikmisi yang diterima penulis selama menjalani perkuliahan di Program Studi Matematika Fakultas Sains dan Teknologi Universitas Jambi.

VI. DAFTAR PUSTAKA

- [1] Awaliyah, R. 2018. *Analisis Clustering Untuk Mengelompokkan Tingkat Kesejahteraan Kabupaten/Kota Berdasarkan Sosial Ekonomi Rumah Tangga di Wilayah Provinsi Sulawesi Selatan*. Skripsi. UIN Alauddin Makassar.
- [2] Badan Pusat Statistik. 2017. *Indikator Kesejahteraan Rakyat 2017 (Welfare Indicators 2017)*. Publikasi No 07330.1713. BPS, Jakarta.
- [3] _____. 2017. *Statistik Indonesia 2018(Statistical Yearbook Of Indonesia 2018)*. Publikasi No 03220.1811. BPS, Jakarta.

- [4] _____ . 2017. *Statistik Kesejahteraan Rakyat 2017 (Welfare Statistic 2017)*. Publikasi No 04210.1718. BPS, Jakarta.
- [5] Gunawan, I. 2016. *Pengantar Statistika Inferensial*. PT. RajaGrafindo Persada, Depok.
- [6] Johnson, R.A dan D.W. Wichern. 2007. *Applied Multivariate Statistical Analysis Sixth Edition*. Prentice Hall International, New Jersey.
- [7] Mattjik, A.A dan I.M Sumertajaya. 2011. *Sidik Peubah Ganda dengan Menggunakan SAS Edisi Pertama*. IPB Press, Bogor.
- [8] Walpole, R.E dan R.H. Myers. 1986. *Ilmu Peluang dan Statistika Untuk Insinyur dan Ilmuan, Terbitan Kedua*. Penerbit ITB, Bandung.