

ANALISIS POPULASI PENDUDUK DAN TINGKAT PENDIDIKAN DI KABUPATEN ACEH JAYA DENGAN METODE *CLUSTER* PAUTAN TUNGGAL

Wiwin Apriani¹, Rahmi Hayati²

^{1,2}Dosen Program Studi Pendidikan Matematika FKIP Universitas Almuslim
Email: wiwina10@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk mengetahui heterogenitas populasi penduduk dan tingkat pendidikan di Kabupaten Aceh Jaya tahun 2020. Data yang digunakan adalah data sekunder hasil dari Sensus Penduduk yang dilakukan oleh Badan Pusat Statistik (BPS). Variabel yang diamati dalam penelitian ini adalah 1) variabel penjelas (X), terdiri atas: (1) populasi penduduk (X_1), berupa data jumlah penduduk, jumlah laki-laki, jumlah perempuan dan jumlah kepala rumah tangga (RT); (2) tingkat pendidikan (X_2), berupa data jumlah belum tamat SD, tamat SD, tamat SMP, tamat SMA, tamat SMK, tamat D1, tamat D3, tamat D4/S1, dan jumlah tamat S2/S3; serta 2) variabel respon (Y), merupakan kecamatan yang terdapat di Kabupaten Aceh Jaya, terdiri dari Teunom, Panga, Krueng Sabee, Setia Bakti, Sampoiniet dan Jaya. Untuk menganalisis kasus tersebut, digunakan analisis *cluster* pautan tunggal, yaitu mengelompokkan setiap obyek pengamatan dalam satu kelompok yang terdiri dari satu anggota. Lalu, menghitung nilai jarak antarobyek dengan metode pautan tunggal. Hasil yang diperoleh menunjukkan bahwa keheterogenan populasi penduduk di Kabupaten Aceh Jaya yang paling besar di Kecamatan Panga dan Jaya dengan jarak 15723,364. Sedangkan pada tingkat pendidikan, Kecamatan Panga dan Jaya terjadi keheterogenan karakteristiknya dengan jarak 9019,637. Hal ini berarti populasi penduduk dan tingkat pendidikan pada Kecamatan Jaya dan Panga memiliki perbedaan karakteristik yang sangat besar.

Kata kunci: *Cluster pautan tunggal; heterogenitas; populasi penduduk; tingkat pendidikan*

PENDAHULUAN

Salah satu metode multivariat yang dapat digunakan untuk mengelompokkan adalah analisis *cluster*, dimana pada kelompok tersebut memiliki karakteristik atau ciri yang relatif sama dan untuk kelompok yang lain memiliki karakteristik atau ciri yang berbeda. Analisis *cluster* merupakan analisis statistik yang dapat digunakan untuk mengelompokkan objek yang diamati. Objek tersebut dikelompokkan ke dalam satu atau beberapa kelompok atau *cluster* berdasarkan kemiripan yang dimilikinya.

Secara umum, suatu objek dimasukkan ke dalam kelompok jika suatu objek memiliki karakteristik yang sama, objek tersebut akan cenderung berkorelasi (berhubungan) dengan objek lainnya pada kelompok tersebut. Pengelompokkan dapat terjadi berdasarkan kuat tidaknya korelasi (hubungan) antarobjek. Secara umum, terdapat dua metode pengelompokan data dalam analisis *cluster*, yaitu metode hierarki dan non hierarki (Bangun, 2016). Lalu, Fathia (2016), menjelaskan bahwa analisis *cluster* hierarki memiliki beberapa metode, yaitu metode pautan tunggal (*single linkage*), pautan lengkap (*complete linkage*), antar pusat (*centroid linkage*), pautan rata-rata (*average linkage*) dan metode ward (*ward's method*), sedangkan pada metode non hierarki terdapat metode *k-means*.

Penelitian terdahulu yang pernah dilakukan adalah penerapan analisis *cluster* dengan metode hierarki untuk klasifikasi kabupaten/kota di Provinsi Maluku berdasarkan indikator indeks pembangunan manusia dilakukan oleh Matdoan (2020). Selanjutnya, Putri (2019) meneliti tentang penerapan analisis *cluster ensemble* untuk mengelompokkan provinsi di Indonesia berdasarkan indikator kesehatan lingkungan. Kemudian, analisis *cluster* menggunakan algoritma *k-means cluster* untuk *clustering* jenis penyakit menular pada puskesmas di Kecamatan Kota Tangerang dilakukan oleh Rachman (2019). Perbandingan hasil analisis *cluster* dengan menggunakan metode *single linkage* dan metode *k-means* diteliti oleh Maria (2016). Analisis *cluster* non hierarki menggunakan metode *k-means* dilakukan oleh Amah (2017). Perbandingan kinerja metode *complete linkage* dan *average linkage* dalam menentukan hasil *cluster* (data produksi palawija Provinsi Kalimantan Timur) dilakukan oleh

Ningsih (2016). Adapun perbedaan penelitian ini dengan penelitian sebelumnya terletak pada perbandingan metode dan data yang digunakan. Pada penelitian ini, hal-hal yang akan dianalisis adalah populasi penduduk dan tingkat pendidikan di Kabupaten Aceh Jaya menggunakan metode *cluster* pautan tunggal. Penelitian ini bertujuan untuk mengetahui heterogenitas populasi penduduk dan tingkat pendidikan di Kabupaten Aceh Jaya tahun 2020.

METODE PENELITIAN

Variabel-variabel yang diamati dalam penelitian ini adalah: 1) variabel penjelas (X), terdiri atas: (1) populasi penduduk (X₁), berupa data jumlah penduduk, jumlah laki-laki, jumlah perempuan dan jumlah kepala rumah tangga (RT); (2) tingkat pendidikan (X₂), berupa data jumlah belum tamat SD, tamat SD, tamat SMP, tamat SMA, tamat SMK, tamat D1, tamat D3, tamat D4/S1, dan jumlah tamat S2/S3; serta 2) variabel respon (Y), merupakan kecamatan yang terdapat di Kabupaten Aceh Jaya, terdiri dari Teunom, Panga, Krueng Sabee, Setia Bakti, Sampoiniet dan Jaya.

Langkah-langkah yang dilakukan dalam analisis *cluster* adalah mengelompokkan setiap obyek pengamatan ke dalam satu kelompok yang terdiri dari satu anggota. Lalu, menghitung nilai jarak antarobyek menggunakan metode pautan tunggal. Metode ini diawali dengan menemukan 2 objek yang memiliki jarak terdekat, lalu digabungkan menjadi satu *cluster*, tahapan ini dilakukan terus menerus sampai terbentuk satu *cluster*. Jarak antar *cluster* didefinisikan sebagai jarak minimal antar semua pengamatan. Matriks jarak *D* antara *cluster* (U, V) dengan *cluster* W sebagai berikut:

$$d_{(u,v)w} = \min (d_{uw}, d_{vw})$$

dimana: $d_{(u,v)w}$ = jarak antara *cluster* (u,v) dengan w
 d_{uw} = jarak antara *cluster* u dengan w
 d_{vw} = jumlah anggota *cluster* w

HASIL DAN PEMBAHASAN

Data yang digunakan adalah data sekunder berupa data jumlah penduduk sebanyak 72,713 jiwa yang diperoleh dari hasil Sensus Penduduk oleh Badan Pusat Statistik (BPS) tahun 2020 di Kabupaten Aceh Jaya. Kabupaten Aceh Jaya secara geografis terletak pada 04°22' - 5°16' Lintang Utara dan 95°02' - 63°03' Bujur Timur. Kabupaten ini terdiri dari 6 kecamatan, 21 mukim dan 172 desa, dengan batas wilayah administrasi yaitu sebelah utara berbatasan dengan Kabupaten Aceh Besar dan Pidie, sebelah selatan berbatasan dengan Kabupaten Aceh Barat dan Samudera Indonesia, sebelah timur berbatasan dengan Kabupaten Pidie dan Aceh Barat, dan sebelah barat berbatasan dengan Samudera Indonesia (Aceh Jaya dalam Angka, 2020). Lalu, penghitungan data menggunakan perangkat lunak SPSS 22.

Populasi Penduduk

Berdasarkan data populasi penduduk Kabupaten Aceh Jaya pada tahun 2020 untuk setiap kecamatan di Kabupaten Aceh Jaya dibentuk menjadi 6 *cluster* yang terdiri dari 1 kecamatan, yaitu Teunom (Tn), Panga (Pn), Krueng Sabe (Ks), Setia Bakti (Sb), Sampoiniet (Sp) dan Jaya (Jy). Lalu, mencari ukuran ketakmiripan *cluster* pada variabel populasi penduduk menggunakan *Software SPSS versi 2*, dan diperoleh hasil ukuran ketakmiripan variabel, sebagai berikut:

Tabel 1. Ukuran Ketakmiripan Variabel Populasi Penduduk

	Tn	Pn	Ks	Sb	Sp	Jy
Tn	0,000	12086,142	3580,964	12028,401	6667,282	3720,320
Pn		0,000	8538,995	207,808	5452,756	15723,296
Ks			0,000	8475,604	3099,377	7198,419
Sb				0,000	5382,408	15653,991
Sp					0,000	10275,931

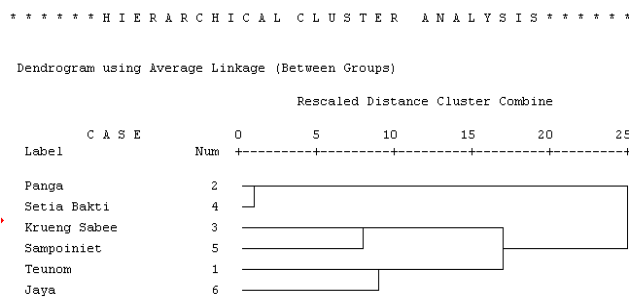
Jy	0,000
----	-------

Pada tabel di atas terlihat bahwa jarak antara Teunom dengan Panga sebesar 12086,142, sedangkan jarak antara Teunom dengan Krueng Sabee sebesar 3580,964. Hal ini menunjukkan bahwa Kecamatan Teunom lebih mirip karakteristiknya (jumlah penduduk, jumlah laki-laki dan perempuan, jumlah kepala RT dan jumlah kematian) dengan Kecamatan Krueng Sabee. Namun, Kecamatan Teunom sangat heterogen dengan Kecamatan Panga. Demikian seterusnya untuk penafsiran variabel yang lain. Semakin besar nilai jarak antara dua variabel, maka semakin heterogen kedua variabel tersebut. Pembentukan *cluster* dengan metode pautan tunggal dilakukan dengan menggabungkan dua variabel yang memiliki jarak terkecil menjadi kelompok yang baru. Proses penggabungan dua variabel atau lebih terlihat pada tabel *Agglomeration Schedule* berikut:

Tabel 2. Agglomeration Schedule Populasi Penduduk

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	4	207,808	0	0	5
2	3	5	3099,377	0	0	4
3	1	6	3720,320	0	0	4
4	1	3	6930,649	3	2	5
5	1	2	10417,699	4	1	0

Adapun tahapan (*stage*) dari table di atas yaitu pada *Stage* 1, terbentuk satu *cluster* antara Panga dengan Setia Bakti, dengan nilai koefisien sebesar 207,808 yang menunjukkan besarnya jarak antara variabel Panga dan Setia Bakti. Proses agglomerasi dimulai dengan jarak antarvariabel yang paling dekat. Lalu, jika berpedoman pada kolom tahap *next stage*, akan diperoleh angka 5. Hal ini berarti langkah *clustering* selanjutnya dilakukan dengan melihat *stage* 5, yaitu terlihat terbentuknya *cluster* baru antara Krueng Sabee dan Sampoiniet dengan nilai koefisien sebesar 10417,699. Lalu, proses *cluster* dilanjutkan hingga kolom tahap selanjutnya menunjukkan nilai tahap 0, yang berarti proses *cluster* berhenti. Visualisasi dari proses agglomerasi ini terlihat pada Dendogram berikut yang memperlihatkan tahapan pembentukan kelompok dalam bentuk dendogram yaitu:



Gambar 1. Dendogram Populasi Penduduk

Berdasarkan gambar di atas terlihat bahwa skala yang digunakan bukan dari koefisien seperti pada tabel *Agglomeration Schedule*, tetapi hasil dari dilakukannya proses skala ulang terhadap batasan antara 0 dan 25. Proses *agglomerasi* dimulai dari skala 0 dengan ketentuan jika sebuah garis dekat dengan angka 0, maka variabel yang terwakili dengan garis tersebut semakin mungkin membentuk sebuah *cluster*. Dendogram berguna untuk menunjukkan anggota *cluster* yang ada, berapapun *cluster* yang akan ditentukan seharusnya terbentuk. Sebagai contoh, jika ditentukan 3 buah *cluster*, maka dari dendogram di atas terlihat *Cluster* I : {Tn, (Ks,Sp)} = Teunom, Krueng Sabee dan Sampoiniet, *Cluster* II : {(Pn,Sb)} = Panga dan Setia Bakti dan *Cluster* III: { Jy } = Jaya.

Berdasarkan Tabel 2 di atas diketahui bahwa jarak antar semua variabel tidak mendekati “no1”. Hal ini berarti semua variabel memiliki keheterogenan dalam hal karakteristiknya (jumlah penduduk, jumlah laki-laki, jumlah perempuan, jumlah kepala RT dan jumlah kematian). Sebagai contoh, jarak antar variabel 1 (Teunom) dengan variabel 2 (Panga) adalah 12.086,142. Sedangkan jarak antar variabel 1 (Teunom) dengan variabel 3 (Krueng Sabee) adalah 3.580,964. Hal ini berarti Teunom lebih mirip karakteristiknya dengan Krueng Sabee dan memiliki karakteristik yang heterogen dengan Panga. Maka, diketahui bahwa kecamatan yang paling besar keheterogenan karakteristiknya yaitu antar Kecamatan Panga dan Jaya dengan jarak terbesar antar kedua variabel tersebut 15.723,296.

Tingkat Pendidikan

Penentuan jarak antar variabel dilakukan dengan cara yang sama seperti pada perhitungan jarak populasi penduduk. Adapun hasil perhitungan jarak antar variabel terlihat sebagai berikut:

Tabel 3. Ukuran Ketakmiripan Antar Variabel Tingkat Pendidikan

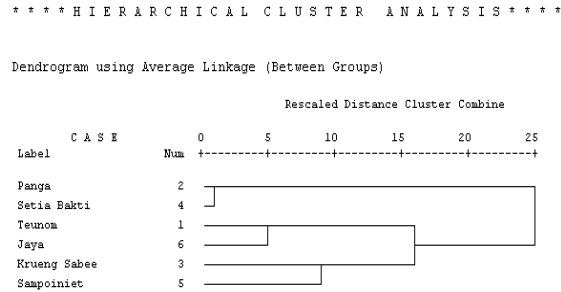
	Tn	Pn	Ks	Sb	Sp	Jy
Tn						1593,828
Pn				7350,977	3984,495	9019,637
Ks	0,000	7476,066	2738,689	543,683	3648,932	4205,437
Sb		0,000	5413,912	5388,834	2592,368	8884,243
Sp			0,000	0,000	3473,554	5502,455
Jy					0,000	0,000

Jarak tiap variabel (jenis kecamatan) yang dihitung dengan jarak *euclidean* untuk tingkat pendidikan ditampilkan dalam tabel di atas. Sebagai contoh, jarak antara Teunom dengan Panga sebesar 7476,066, sedangkan jarak antara Teunom dengan Krueng Sabee sebesar 2738,689. Hal ini menunjukkan bahwa Kecamatan Teunom lebih mirip karakteristiknya (jumlah penduduk, partisipasi sekolah dan tingkat pendidikan formal) dengan Kecamatan Krueng Sabee, tetapi heterogen dengan Kecamatan Panga. Demikian seterusnya untuk penafsiran variabel lain. Proses penggabungan dua variabel atau lebih dapat dilihat pada Tabel *Agglomeration Schedule* berikut ini:

Tabel 4. Agglomeration Schedule Tingkat Pendidikan

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2	4	543,683	0	0	5
2	1	6	1593,828	0	0	4
3	3	5	2592,368	0	0	4
4	1	3	4107,769	2	3	5
5	1	2	6332,019	4	1	0

Dari tabel di atas terlihat beberapa tahapan yaitu pada *Stage 1*, terbentuk satu *cluster* antara Panga dengan Setia Bakti, dengan nilai koefisien sebesar 543,683 yang menunjukkan besarnya jarak antara variabel Panga dan Setia Bakti. Proses *Agglomerasi* dimulai dengan jarak antar variabel yang paling dekat. Lalu, jika didasarkan pada kolom *next stage* diperoleh angka 5. Hal ini berarti langkah *clustering* selanjutnya dilakukan dengan melihat *stage 5*. Proses dari *cluster* diteruskan hingga diperoleh pada kolom tahap selanjutnya menunjukkan 0, dimana proses *cluster* akan berhenti. Visualisasi dari proses agglomerasi ini dapat dilihat pada Dendogram gambar berikut ini:



Gambar 2. Dendrogram Tingkat Pendidikan

Berdasarkan gambar di atas diperoleh bahwa skala yang digunakan bukan berasal dari koefisien seperti pada tabel *Agglomeration Schedule*, tetapi hasil dari dilakukannya proses skala ulang terhadap batasan antara 0 dan 25. Dendrogram berguna untuk menunjukkan anggota *cluster* yang ada, berapapun *cluster* yang akan ditentukan seharusnya terbentuk. Sebagai contoh, jika ditentukan 3 buah *cluster*, maka dari dendrogram di atas terlihat bahwa *Cluster I* : {Tn, Jy} = Teunom, Jaya, *Cluster II* : {(Pn,Sb)} = Panga, Setia Bakti, dan *Cluster III*: {Kr, Sp} = Krueng Sabee, Sampoiniet.

Lalu, dari tabel 3 diperoleh jarak antar variabel 1 (Teunom) dan variabel 4 (Setia Bakti) adalah 7350,977. Sedangkan jarak antar variabel 2 (Panga) dan variabel 4 (Setia Bakti) adalah 543,683. Hal ini berarti Kecamatan Setia Bakti lebih mirip karakteristiknya dengan Kecamatan Teunom dan memiliki karakteristik heterogen dengan Kecamatan Panga. Maka, kecamatan yang paling besar keheterogenan karakteristiknya yaitu antar Kecamatan Panga dan Jaya dengan jarak antar kedua variabel 9019,637.

SIMPULAN

Berdasarkan hasil dan pembahasan keheterogenan dari dua buah sektor yang diamati, diperoleh bahwa keheterogenan paling besar terjadi pada saat nilai jarak *euclidean* yang dihasilkan paling besar. Maka, disimpulkan bahwa populasi penduduk di Kabupaten Aceh Jaya yang paling besar keheterogenannya yaitu di Kecamatan Panga dan Jaya dengan jarak *euclidean* 15723,296. Sedangkan untuk tingkat pendidikan paling besar terjadi di Kecamatan Panga dan Jaya dengan jarak *euclidean* sebesar 9019,637. Sehingga, populasi penduduk dan tingkat pendidikan pada Kecamatan Jaya dan Panga memiliki perbedaan karakteristik yang sangat besar.

REFERENSI

- Amah, N., et. al. 2017. *Analisis Cluster Non-Hirarki dengan Metode K-Modes pada Mahasiswa Prodi. Statistika Angkatan 2015 FMIPA Universitas Mulawarman*. Jurnal Eksponen, Vol. 8 No. 1 Hal. 9-16.
- Bangun, R.H.B. 2016. *Analisis Klaster Non Hierarki dalam Pengelompokan Kabupaten di Sumatera Utara Berdasarkan Faktor Produksi Padi*. Agrica (Jurnal Agribisnis Sumatera Utara), Vol. 4 No. 1 Hal. 54-61.
- Fathia, A.N., Rahmawati, R., Tarno. 2016. *Analisis Klaster Kecamatan di Kabupaten Semarang Berdasarkan Potensi Desa Menggunakan Metode Ward dan Single Linkage*. Jurnal Gaussian, Vol. 5 No. 4 Hal. 801-810.

- Maria, G., Wahyuningsih, S. 2016. *Perbandingan Hasil Analisis Cluster dengan Metode Single Linkage dan Metode C-Means (Studi Kasus: Data Tingkat Kualitas Udara Ambien pada Perusahaan Perkebunan di Kab. Kutai Barat Tahun 2014)*. Jurnal Eksponensial, Vol. 7 No. 1 Hal. 9-16.
- Matdoan, M.Y., Van Delsen, M.S.N. 2020. *Penerapan Analisis Cluster dengan Metode Hierarki untuk Klasifikasi Kabupaten/Kota di Provinsi Maluku Berdasarkan Indikator Indeks Pembangunan Manusia*. Statmat (Jurnal Stat. dan Mat., Vol. 2 No. 2 Hal. 123-130.
- Ningsih, S., Wahyuningsih, S., Nasution, Y. N. 2016. *Perbandingan Kinerja Metode Complete Linkage dan Average Linkage dalam Menentukan Hasil Analisis Cluster (Studi Kasus: Produksi Palawija Provinsi Kalimantan Timur 2014/2015)*. Prosiding Seminar Sains dan Teknologi FMIPA Unmul, Vol. 1 No. 1 Hal. 46-50.
- Putri, Y.D., Yozza, H. 2019. *Penerapan Analisis Cluster Ensemble untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Indikator Kesehatan Lingkungan*. Jurnal Mat. UNAND, Vol. VIII No. 1 Hal. 323-330.
- Rachman, A., Aden., Rusdiana, Y. 2019. *Analisis Cluster Menggunakan Algoritma K-Means Cluster untuk Culstering Jenis Penyakit Menular pada Puskesmas di Kecamatan Kota Tangerang*. Jurnal Saintika UNPAM, Vol. 2 No. 1 Hal. 15-27.