

Analisis *Cluster Longitudinal* pada Pembangunan Manusia di Sulawesi Selatan Berbasis Gender

Khalilah Nurfadilah

Universitas Islam Negeri Alauddin Makassar, khalilahnurfadilah1202@gmail.com

ABSTRAK Disparitas capaian IPG dan IDG masih relatif tinggi antar kabupaten/kota di Sulawesi Selatan. Masalah lain yang muncul pada pembangunan manusia di Sulawesi Selatan adalah capaian IPG yang tinggi namun tingkat IPM rendah dan sebaliknya capaian IPM yang tinggi namun IPG rendah. Perlu formulasi khusus untuk menetapkan pendekatan pembangunan yang berorientasi pada kesetaraan dan keadilan gender. Salah satu solusinya adalah dengan mengklasifikasikan kabupaten/kota berdasarkan capaian IPM, IPG, dan IDG sehingga intervensi pemerintah tepat sasaran. Klasifikasi dilakukan dengan metode analisis *Cluster Longitudinal* dengan metode *K-Means* jarak *Manhattan*. Ruang lingkup penelitian terdiri dari 24 kabupaten/Kota di Sulawesi Selatan pada periode waktu 2010-2020. Hasil dari *Calinski & Harabasz Criterion* menunjukkan cluster optimum sebanyak 6 cluster. Pembahasan hasil menunjukkan bahwa daerah perkotaan di Sulawesi Selatan relatif lebih sukses dalam pembangunan manusia berbasis gender, dapat dibuktikan dengan cluster E yang beranggotakan kota Makassar, Parepare, dan Palopo dengan kategori IPM perempuan, IPG, dan IDG yang relatif tinggi dibandingkan cluster lain. Sedangkan kabupaten pada cluster C, walaupun IDG-nya relatif tinggi namun capaian IPM perempuan dan IPG relatif rendah.

Kata Kunci: *Cluster Longitudinal, IPG, IDG, jarak Manhattan, K-Means*

1. PENDAHULUAN

Menurut data Badan Pusat Statistik (BPS), sejak tahun 2017 capaian Indeks Pembangunan Manusia (IPM) di Sulawesi Selatan telah masuk kategori tinggi yang menunjukkan keseriusan pemerintah terhadap pembangunan yang tidak hanya dititikberatkan terhadap infrastruktur saja, tetapi juga manusianya. Namun, keberhasilan pembangunan manusia ini belum sepenuhnya berhasil, jika dipilah berdasarkan jenis kelamin, capaian IPM masih menunjukkan kesenjangan pada perempuan.

Menurut data BPS pada tahun 2020, capaian IPM laki-laki di Sulawesi Selatan telah masuk kategori tinggi yaitu 75,90 sedangkan capaian IPM perempuan Sulawesi Selatan yaitu 70,48 masuk juga kategori tinggi, namun selisihnya masih relatif tinggi. Padahal, dalam

Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020-2024 telah mengarusutamakan *Sustainable Development Goals* (SDGs) dimana kesetaraan gender menjadi salah satu tujuan utama yang harus diwujudkan pada tahun 2030[1].

Alat ukur yang digunakan untuk mengevaluasi hasil pembangunan yang berperspektif gender antara lain adalah Indeks Pemberdayaan Gender (IDG) dan Indeks Pembangunan Gender (IPG). Menilik dari alat ukur tersebut, capaian IPG dan IDG Sulawesi Selatan pada tahun 2020 relatif tinggi, yaitu masing-masing sebesar 92,86 dan 76,32 yang melebihi capaian IPG dan IDG nasional[2]. Namun capaian ini belum menyeluruh sampai ke level kabupaten/kota. Dari 24 kabupaten/kota di Sulawesi Selatan, baru 11 kabupaten/kota yang IPG diatas capaian provinsi, bahkan hanya 4 kabupaten/kota yang capaian IDG diatas provinsi.

Disparitas capaian IPG dan IDG masih relatif tinggi antar kabupaten/kota di Sulawesi Selatan. Masalah lain yang muncul pada pembangunan manusia di Sulawesi Selatan adalah capaian IPG yang tinggi namun tingkat IPM rendah dan sebaliknya capaian IPM yang tinggi namun IPG rendah[3].

Pemangku kebijakan perlu formulasi khusus untuk menetapkan pendekatan pembangunan yang berorientasi pada kesetaraan dan keadilan gender. Salah satu solusinya adalah dengan mengklasifikasikan kabupaten/kota berdasarkan capaian IPM, IPG, dan IDG sehingga intervensi pemerintah tepat sasaran. Klasifikasi dilakukan dengan metode analisis *Cluster Longitudinal* yang lebih kompleks dan terukur dibanding analisis *cluster* data *cross section*[4].

Analisis *cluster longitudinal* terdiri dari berbagai macam metode, salah satunya metode *K-Means* dengan jarak *Manhattan* yang lebih optimum dibandingkan jarak *Eclidean*,

Maximum, Frechet dan *Dynamic Time Warping*[5].

2. TINJAUAN PUSTAKA

Indeks Pembangunan Gender

Indeks pembangunan gender adalah perbandingan antara IPM laki-laki dan IPM perempuan ditinjau dari aspek kualitas dimensi pendidikan, kesehatan, dan ekonomi. Dimensi pendidikan menggunakan harapan lama sekolah dan rata-rata lama sekolah, dimensi kesehatan menggunakan angka harapan hidup, dan dimensi ekonomi menggunakan pengeluaran perkapita. Angka IPG yang menghampiri 100 menunjukkan bahwa pencapaian pembangunan perempuan hamper sama dengan laki-laki. Secara persamaan, IPG dapat dirumuskan sebagai berikut[6] :

$$IPG = \frac{IPM_{laki-laki}}{IPM_{perempuan}} \quad (2.1)$$

Indeks Pemberdayaan Gender

Indeks pemberdayaan perempuan digunakan untuk mengukur partisipasi aktif perempuan di bidang ekonomi, politik, dan manajerial. Tiga indikator yang digunakan yaitu distribusi pendapatan perempuan terhadap pendapatan kerja, keterwakilan perempuan di parlemen, dan keterlibatan perempuan dalam pengambilan keputusan melalui indikator perempuan sebagai tenaga manajerial, professional, administrasi, dan teknisi[7].

Analisis Cluster

Analisis cluster bertujuan untuk menggabungkan beberapa objek kedalam kelompok-kelompok berdasarkan sifat kemiripan atau sifat ketidakmiripan antar objek[8]. Ada dua metode dalam analisis cluster, yaitu :

1. Metode berhierarki : dimulai dengan mengelompokkan dua atau lebih objek yang mempunyai kesamaan paling dekat. Kemudian dilanjutkan ke objek lain yang mempunyai kedekatan kedua, demikian seterusnya sehingga klaster terbentuk semacam pohon berhierarki antara objek.

Untuk memperjelas proses hierarki digunakan dendogram.

2. Metode tak berhierarki : apabila banyaknya kelompok yang akan dibentuk sudah diketahui. Metode *K-means* merupakan salah satu contohnya. Pada metode *K-means*, besaran k ditentukan terlebih dahulu.

K-Means

Metode *K-means* berupaya mempartisi data yang ada kedalam bentuk satu atau lebih klaster, sehingga data dengan karakteristik yang sama diklaster dalam satu kelompok yang sama. Berikut adalah algoritma dari *K-means* :

1. Penentuan jumlah kelompok dan alokasikan data kedalam kelompok secara acak.
2. Tentukan centroid setiap kelompok dengan mengambil rata-rata semua nilai dari setiap fiturnya. Persemaaan *centroid* fitur ke- i sebagai berikut :

$$C^1 = \frac{1}{M} \sum_{j=1}^M X_j \quad (2.2)$$

Dimana M adalah jumlah data dalam kelompok, i adalah fitur ke- i , dan p adalah dimensi data. Persamaan diatas dilakukan sebanyak p dimensi $i = 1$ sampai dengan $i = p$.

3. Masing-masing data dialokasikan ke *centroid* terdekat. Salah satu metode pengukuran yang digunakan adalah jarak *Euclidean* dengan persamaan berikut :

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (2.3)$$

Pengalokasian kembali data kedalam masing-masing kelompok pada metode *K-means* didasarkan pada perbandingan jarak antara data dengan *centroid* setiap kelompok yang ada. Data dialokasikan kembali ke kelompok yang mempunyai *centroid* dengan jarak terdekat dari data tersebut. Alokasi data ini ditentukan melalui persamaan berikut :

$$a_{i1} = \begin{cases} 1, & d = \min\{D(X_i, C_i)\} \\ 0, & \text{lainnya} \end{cases} \quad (2.4)$$

Dimana a_{i1} merupakan nilai keanggotaan titik x_1 ke pusat kelompok C_i , d adalah jarak terpendek dari data x_1 ke K kelompok setelah dibandingkan, dan C_i adalah centroid ke-1. Fungsi objektif yang digunakan untuk metode *K-means* ditentukan berdasarkan jarak dan nilai keanggotaan data dalam kelompok. Persamaan fungsi objektif sebagai berikut :

$$J = \sum_{i=0}^n \sum_{i=1}^k a_{ic} D(x_i, c_i)^2 \quad (2.5)$$

Dimana n adalah jumlah data, k adalah jumlah kelompok, a_{ic} adalah nilai keanggotaan titik data x_i ke kelompok c_i yang diikuti, a_{ic} mempunyai nilai 1 jika data adalah anggota kelompok dan 0 jika data bukan anggota.

Analisis Cluster *K-Means* Jarak Manhattan

Analisis klaster pada data longitudinal berbeda dengan analisis klaster pada data *cross section*, dalam klasterisasi ini algoritma yang digunakan harus dapat mendukung struktur data panel. Metode *K-Means* dapat digunakan dalam pengelompokan data panel dengan memodifikasi ukuran jarak yang disesuaikan dengan struktur data panel. Namun, ada beberapa pendekatan yang dapat digunakan untuk *klustering* data panel. Jarak *Manhattan* memiliki kinerja yang paling optimum dibanding jarak *Maximum*, *Frechet*, *Euclidean*, dan *Dynamic Time Wrapping*. Jarak *Manhattan* didefinisikan pada persamaan berikut dengan $p=1$:

$$Dist(y_{1..}, y_{2..}) = \sum_{j=1}^T \sum_X^M [y_{1jX} - y_{2jX}] \quad (2.6)$$

Jarak maksimum diperoleh dengan limit $p \rightarrow +\infty$, dengan persamaan sebagai berikut :

$$Dist(y_{1..}, y_{2..}) = \left\{ \sum_{j=1}^T \sum_X^M [y_{1jX} - y_{2jX}]^p \right\}^{\frac{1}{p}}, p \rightarrow \infty \quad (2.7)$$

Pada penelitian ini jumlah klaster tidak ditentukan. Untuk melihat jumlah klaster yang optimal digunakan kriteria nilai Calinski dan

Harabasz. Kriteria Calinski dan Harabasz diberikan pada persamaan berikut[9] :

$$C(k) = \frac{Trace(B)}{Trace(W)} \frac{n-k}{k-1} \quad (2.8)$$

3. METODE

Data yang digunakan adalah data sekunder yang didapatkan dari website BPS dengan objek pengamatan 24 kabupaten/kota di Sulawesi Selatan pada periode tahun 2010-2020. Adapun variabel-variabel yang diperhatikan adalah IPM Perempuan (X_1), IPG (X_2), dan IDG (X_3).

Prosedur Penelitian

Prosedur penelitian ini dijelaskan sebagai berikut:

1. Menginputkan data dalam bentuk matriks.
2. Proses analisis *Cluster Longitudinal* menggunakan metode *K-Means*.
3. Menentukan jumlah *cluster* optimum menggunakan kriteria nilai Calinski dan Harabasz.
4. Menentukan kemiripan objek menggunakan jarak *Manhattan*.
5. Melakukan interpretasi dengan melihat nilai centroid untuk menentukan kategori masing – masing *cluster*.
6. Mengklasifikasikan kelompok-kelompok yang terbentuk berdasarkan kategori dengan metode distribusi frekuensi numerik yang disusun berdasarkan interval.

4. PEMBAHASAN

Profil Data

Terdapat 24 kabupaten/kota di Sulawesi Selatan sebagai objek pengamatan pada periode waktu 2010-2020, yang akan diklasterkan dengan analisis *cluster longitudinal*.

Cluster Optimum

Jumlah kelompok tidak ditentukan, sehingga untuk menentukan jumlah kelompok optimum digunakan nilai kriteria Calinski dan Harabasz. Pada Tabel 4.1 terlihat nilai optimum terbentuk pada *cluster* 6 dengan nilai 18,675.

Tabel 4.1 Calinski & Harabsz *Criterion*

Jumlah Cluster (k)	Nilai CH
2	17,924
3	15,058
4	17,210
5	16,925
6	18,675
7	17,238
8	15,761
9	15,568
10	15,152

Hasil dari *clustering* membentuk 6 kelompok dengan karakteristik yang homogen antaranggota dan heterogen antarkelompok. Rincian anggota dan jumlah kelompok yang terbentuk dapat dilihat pada Tabel 4.2 yang menunjukkan *Cluster A* yang memiliki anggota terbanyak dengan 7 anggota dan *Cluster F* yang hanya memiliki 1 anggota.

Tabel 4.2 Hasil Analisis *Cluster*

Cluster	Kabupaten/Kota	Jumlah Anggota
A	Kepulauan Selayar, Jeneponto, Maros, Pangkep, Bone, Wajo, Luwu	7
B	Bulukumba, Sinjai, Barru, Soppeng, Pinrang, Enrekang	6
C	Takalar, Gowa, Tana Toraja, Toraja Utara	4
D	Sidenreng Rappang, Luwu Utara, Luwu Timur	3
E	Makassar, Parepare, Palopo	3
F	Bantaeng	1

Klasifikasi *Cluster*

Clustering kabupaten/kota di Sulawesi Selatan pada periode waktu 2010-2020 yang menghasilkan 6 *cluster* dengan *mean* masing-masing *cluster* seperti pada Tabel 4.3 yang selanjutnya akan diklasifikasikan menggunakan tabel distribusi.

Tabel 4.3 *Mean* hasil *Clustering*

<i>Mean</i>	X_1	X_2	X_3
Cluster A	62,55	90,06	61,86

Cluster B	66,20	96,29	63,43
Cluster C	60,24	85,82	66,71
Cluster D	65,17	89,47	46,68
Cluster E	75,62	95,60	68,47
Cluster F	63,86	96,00	77,55

Setelah diperoleh *mean* dari masing-masing *cluster*, selanjutnya menentukan jumlah kelas, panjang kelas, dan rentang untuk menentukan batas-batas kategori pada tabel distribusi frekuensi numerik. Hasilnya dapat dilihat pada Tabel 4.4.

Tabel 4.4 Distribusi frekuensi numerik berdasarkan kategori

Kategori	X_1	X_2	X_3
Rendah	$M < 64,37$	$M < 88,31$	$M < 55,97$
Sedang	$64,37 \leq M \leq 69,49$	$88,31 \leq M \leq 91,80$	$55,97 \leq M \leq 66,26$
Tinggi	$M > 69,49$	$M > 91,80$	$M > 66,26$

Dari Tabel 4.4 setiap *cluster* dapat diklasifikasikan berdasarkan variabel X_1 , X_2 , dan X_3 dengan tiga kategori yaitu tinggi, rendah, dan sedang seperti yang ditunjukkan pada Tabel 4.5.

Tabel 4.5 Klasifikasi kategori setiap *Cluster*

Kategori	X_1	X_2	X_3
Cluster A	Rendah	Sedang	Sedang
Cluster B	Sedang	Tinggi	Sedang
Cluster C	Rendah	Rendah	Tinggi
Cluster D	Sedang	Sedang	Rendah
Cluster E	Tinggi	Tinggi	Tinggi
Cluster F	Rendah	Tinggi	Tinggi

Tabel 4.5 menyimpulkan bahwa *cluster E* merupakan *cluster* yang anggotanya berklasifikasi tinggi, dibuktikan dengan setiap variabelnya berkategori tinggi. Dapat diinterpretasikan bahwa *cluster E*, pembangunan manusia yang berorientasi gender relatif tinggi. Berbeda dengan *cluster C* yang menunjukkan variabel X_1 dan X_2 berkategori rendah walaupun X_3 tinggi, yang menunjukkan bahwa pembangunan manusianya rendah dan ketimpangan gender tinggi.

5. KESIMPULAN

Pembahasan hasil menunjukkan bahwa daerah perkotaan di Sulawesi Selatan relatif lebih sukses dalam pembangunan manusia berbasis gender, dapat dibuktikan dengan *cluster* E yang beranggotakan kota Makassar, Parepare, dan Palopo dengan kategori IPM perempuan, IPG, dan IDG yang relatif tinggi dibandingkan cluster lain. Sedangkan kabupaten pada *cluster* C, walaupun IDG-nya relatif tinggi namun capaian IPM perempuan dan IPG relatif rendah.

Pada kasus di *cluster* F yang beranggotakan kabupaten Bantaeng, capaian IPG relatif tinggi namun IPM-nya rendah, diharapkan pembangunan berbasis gender yang berkelanjutan dengan kategori sama-sama tinggi.

6. DAFTAR PUSTAKA

- [1] Kementerian Perencanaan Pembangunan Nasional Republik Indonesia/Badan Perencanaan Pembangunan Nasional (Kementerian PPN/Bappenas), "Rencana Pembangunan Jangka Menengah Nasional 2020-2024", 2019.
- [2] Badan Pusat Statistik (BPS), "Indeks Pembangunan Gender (IPG) 2020", 2021.
- [3] Badan Pusat Statistik (BPS) Sulawesi Selatan, "Indeks Pembangunan Gender Provinsi Sulawesi Selatan 2019", 2020.
- [4] A.S. Awaluddin and I. Taufik, "Analisis *Cluster Data Longitudinal* pada pengelompokan daerah berdasarkan indikator IPM di Jawa Barat", *Prosiding Seminar Nasional Metode Kuantitatif*, 2017.
- [5] A.S. Cahyani Sugiono, Indahwati, and I.M. Sumertajaya, "Comparison of multiple distances on panel data in non-hierarchical clustering method", *International Journal of Scientific and Research Publication*, vol 10:7, 2020.
- [6] Kementerian Pemberdayaan Perempuan dan Perlindungan Anak Republik Indonesia, "Pembangunan Manusia Berbasis Gender", 2020.
- [7] N.S. Wisnujati, "Penyusunan Indeks Pemberdayaan Gender dan Indeks Pembangunan Kabupaten Bojonegoro", *Jurnal Ilmu Ekonomi Universitas Wijaya Kusuma Surabaya*, Vol. 20:2, 2020.
- [8] I.M. Sumertajaya, "Analisis Gerombol Menggunakan Metode *Two Steps Cluster*", *Forum Statistika dan Komputasi*, Vol.12:1, 2007.
- [9] C. Genolini, X. Alacoque, M. Sentenac, and C. Arnaud, "KML and KML3D:R Packages to Cluster Longitudinal Data", *JSTATSOF*, Vol.65:4, 2015.