

Analisis *Clustering Time Series* untuk Pengelompokan Provinsi di Indonesia Berdasarkan Indeks Pembangunan Manusia Jenis Kelamin Perempuan

DWI AGUSTIN NURIANI SIRODJ^{1,2}, I MADE SUMERTAJAYA², ANANG KURNIA²

¹⁾ Program Studi Statistika Fakultas MIPA Universitas Islam Bandung, Indonesia

²⁾ Departemen Statistika, FMIPA, IPB University, Indonesia

e-mail: dwi.agustinnuriani@apps.ipb.ac.id

ABSTRAK

Indeks Pembangunan Manusia (IPM) mencerminkan bagaimana kualitas dari pembangunan suatu wilayah tertentu. Selain adanya ketimpangan nilai IPM antar wilayah provinsi di Indonesia, jika dilihat dari sudut pandang gender, maka kesenjangan IPM laki-laki dan perempuan pun tidak bisa dihindari. Peningkatan pertumbuhan pembangunan di setiap wilayah tentu harus mendorong peningkatan kesetaraan gender pula, dalam hal ini kesenjangan pembangunan antara laki-laki dan perempuan harus mampu diminimalisir sehingga penting untuk melihat bagaimana kondisi IPM perempuan perwilayah provinsi di Indonesia agar dapat dilakukan langkah-langkah intervensi untuk meminimalisir isu ketimpangan yang harapannya dapat mendorong indeks pembangunan di wilayah tersebut. Metode analisis yang digunakan untuk mengelompokkan daerah berdasarkan nilai IPM perempuan adalah *Clustering time series*. Hasil analisis memperlihatkan metode *clustering time series* dengan menggunakan jarak *dynamic time-warping* (DTW) menghasilkan dua kelompok yaitu kelompok 1 (daerah dengan IPM perempuan rendah): daerah Papua dan kelompok 2 (daerah dengan IPM perempuan tinggi): daerah lain selain Papua. Pengelompokan yang dibentuk menghasilkan nilai koefisien *Silhouette* sebesar 0,74. Nilai tersebut menandakan bahwa kelompok yang dibentuk berada dalam kategori kuat dalam artian bahwa dua kelompok tersebut mempunyai karakteristik yang jelas berbeda sehingga metode pengelompokan dengan jarak DTW dapat digunakan dalam pengelompokan provinsi-provinsi di Indonesia berdasarkan nilai IPM Perempuan.

Kata Kunci: IPM Perempuan, Dynamic Time Warping, Clustering time series.

ABSTRACT

The Human Development Index (HDI) reflects the quality of development in a particular region. In addition to the inequality of HDI values between provinces in Indonesia, when viewed from a gender perspective, the gap between the HDI of men and women is inevitable. Increased development growth in each region must certainly encourage an increase in gender equality as well; in this case, the development gap between men and women must be able to be minimized, so it is important to see how the condition of the women's HDI per region in Indonesia so that intervention steps can be taken to minimize the issue of inequality. The analysis method used in this paper is Time Series Clustering. The analysis results show that the time series clustering method using dynamic time-warping (DTW) distance produces two groups: group 1 (regions with low female HDI): Papua region and group (2 regions with high female HDI): all provinces except Papua. The grouping formed produced a *Silhouette coefficient* value of 0.74. This value indicates that the groups formed are in a strong category, so the clustering method with DTW distance can be used in grouping provinces in Indonesia based on the value of Women's HDI.

Kata Kunci: Women's HDI, Dynamic Time Warping, Clustering time series.

1. PENDAHULUAN

Kondisi geografis Indonesia yang terdiri dari sekitar 17.504 pulau membuat kondisi pembangunan wilayah belum merata. Masih adanya ketimpangan dalam hal pembangunan manusia antar provinsi di Indonesia mengakibatkan terjadinya kesenjangan terhadap mutu manusia itu sendiri dan tentunya menjadi masalah yang memerlukan solusi. Salah satu ukuran yang dapat melihat kualitas dari pembangunan suatu wilayah tertentu adalah indeks pembangunan manusia (IPM) (Prastyo & Ilfana, 2022). IPM melihat kondisi pembangunan manusia di suatu wilayah berdasarkan empat dimensi yaitu *Income* (pendapatan per kapita), *Expectation of life* (angka harapan hidup), *Years of scholly* (pendidikan) dan *Elitaration rate* (tingkat buta huruf) (Purba, et al., 2022). Kesenjangan nilai IPM antarwilayah pasti akan memengaruhi capaian IPM secara keseluruhan. Jika provinsi atau kabupaten/kota dengan nilai IPM rendah tidak didorong untuk mempercepat peningkatan kualitas manusia melalui berbagai indikator IPM, capaian daerah lain yang memiliki nilai IPM tinggi menjadi kurang bermakna. (Wahyudi, 2021).

Selain adanya ketimpangan nilai IPM antar wilayah provinsi di Indonesia, jika dilihat dari sudut pandang gender, maka kesenjangan IPM laki-laki dan perempuan pun tidak bisa dihindari. Gender tidak selalu merujuk pada jenis kelamin, akan tetapi gender merujuk pada hubungan antara laki-laki dan perempuan, anak laki-laki dan anak perempuan, dan bagaimana hubungan sosial ini dikonstruksikan yang mana sifatnya dinamis dan dapat berubah antar waktu (Arifin, 2018).

Dilansir dari (Sebayang, 2022) memperlihatkan bahwa selalu terdapat selisih dari IPM laki-laki dan IPM perempuan. Pada tahun 2020, IPM laki-laki mencapai 75,98 atau masuk ke dalam kategori tinggi, sedangkan IPM perempuan masih dalam kategori sedang. Namun, angka harapan hidup perempuan lebih panjang daripada laki-laki dengan angka harapan hidup perempuan 73,5 tahun, sedangkan angka harapan hidup laki-laki 69,6 tahun. Peningkatan pertumbuhan pembangunan di setiap wilayah tentu harus mendorong peningkatan kesetaraan gender pula, dalam hal ini kesenjangan pembangunan antara laki-laki dan perempuan harus mampu diminimalisir sehingga penting untuk melihat bagaimana kondisi IPM perempuan perwilayah provinsi di Indonesia agar dapat dilakukan langkah-langkah intervensi untuk meminimalisir isu ketimpangan yang diharapkan nya dapat mendorong indeks pembangunan di wilayah tersebut.

Kajian mengenai pengelompokan daerah berdasarkan IPM sudah dilakukan oleh (Purba, et al., 2022) yang menggolongkan kabupaten/ kota di Sumatera utara menggunakan metode *K-means clustering*. Hasil penelitian tersebut memperlihatkan bahwa banyak nya kelompok/ *cluster* yang terbentuk adalah sebanyak tiga *cluster* yakni *cluster* dengan tingkat IPM tinggi yang terdiri dari 11 daerah, *cluster* dengan tingkat IPM sedang terdiri dari 18 daerah, dan *cluster* dengan tingkat IPM rendah terdiri dari 4 daerah. Selanjutnya (Ambarwati, 2019) menggunakan metode *latent class cluster* dalam mengelompokkan kabupaten/kota di daerah Jawa timur berdasarkan IPM tahun 2017. Hasil penelitian membentuk dua kelompok; kelompok pertama terdiri dari kabupaten atau kota dengan pembangunan manusia rendah, sedangkan kelompok kedua terdiri dari kabupaten atau kota dengan pembangunan manusia tinggi. Beberapa penelitian sebelumnya baru berfokus pada pengelompokan IPM secara keseluruhan belum spesifik melihat ketimpangan IPM berdasarkan gender. Selain itu, metode yang digunakan belum mengkombinasikan aspek data panel dengan *series* waktu tertentu.

Penambahan aspek *series* waktu dalam proses pengelompokan dapat memberikan informasi yang lebih kaya dibandingkan dengan metode pengelompokan klasik umumnya. (Li & Prakash, 2011) mengemukakan bahwa metode pengelompokan klasik seperti *k-means* tidak dapat mengatasi *lags in sequences* yang mengakibatkan hasil pengelompokan menjadi buruk. Hal ini salah satunya disebabkan karena kejadian saat ini dapat dipengaruhi oleh kejadian masa lalu, sehingga aspek *series* waktu dirasa perlu untuk diperhatikan. Selain itu, melalui menambahkan aspek *series* waktu dapat mengidentifikasi dan meringkas pola dan hubungan dari data (Paparrizos & Gravano, 2015).

Analisis *cluster time series* merupakan metode pengelompokan yang memerlukan data *time series*. Elemen *time series* terpenting yang harus diperhatikan adalah penggunaan ukuran jarak, algoritma *clustering* yang digunakan, dan validasi hasil *cluster* (Aghabozorgi, et al., 2015). Prosedur dari *clustering time series* berbeda dengan data statis, karena menggunakan data *time series* yang diperoleh dari satu titik waktu berbeda dalam interval waktu yang sama dan deretan data tersebut diasumsikan saling berhubungan (Liao, 2005). Terdapat dua kategori utama dalam melakukan *clustering time series* yaitu; (1) *Whole clustering* dimana pengelompokan dilakukan

pada *series* waktu pada level individu, pola-pola yang mirip akan dikelompokkan dalam *cluster* yang sama; dan (2) *Subsequence clustering* dimana pengelompokan dilakukan berdasarkan *time window* (Wang, et al., 2006).

Penelitian ini bertujuan untuk menerapkan metode *clustering time series* dalam mengelompokkan provinsi di Indonesia berdasarkan indeks pembangunan manusia menurut jenis kelamin perempuan. Hasil pengelompokan provinsi diharapkan dapat menjadi rujukan dalam penentuan daerah prioritas percepatan pembangunan nasional dengan memaksimalkan potensi perempuan di daerah tersebut.

2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder mengenai Indeks Pembangunan Manusia berdasarkan jenis kelamin perempuan pada 34 Provinsi di Indonesia tahun 2013-2021 yang diambil dari Badan Pusat Statistik pada laman <https://www.bps.go.id/indicator/40/462/1/indeks-pembangunan-manusia-ipm-menurut-jenis-kelamin.html>. Metode pengelompokan yang dilakukan adalah analisis *clustering time series* dengan menggunakan jarak Dynamic Time Warping (DTW).

Analisis cluster adalah analisis yang mengelompokkan objek berdasarkan kemiripan karakteristik. Pada analisis cluster tidak memiliki peubah respon, sehingga data dari peubah prediktor akan digunakan untuk mengelompokkan objek yang ada. (Mattjik & Sumertajaya, 2011). Pengelompokan objek pada data deret waktu sangat menguntungkan karena akan membentuk pola yang menarik dalam dataset. Analisis pengelompokan deret waktu terbagi menjadi tiga kategori (Liao, 2005) yaitu; (1) pendekatan berbasis data mentah (*raw-data-based*), cara kerjanya menggunakan data mentah untuk menghitung jarak antar deret waktu; (2) pendekatan yang menyarikan data mentah dengan cara menghilangkan pengaruh data noise dan mengurangi dimensi data (*featured-based*), kemudian melakukan perhitungan jarak dan proses pengelompokan; dan (3) metode berbasis model yang bekerja untuk memperbaiki model original dari himpunan data (*model-based*).

Metode cluster hierarki merupakan metode pengelompokkan berdasarkan kemiripan karakteristiknya sementara banyaknya cluster yang diperoleh belum diketahui. Terdapat dua teknik pada metode cluster hierarki yaitu, teknik penggabungan (*agglomerative*) dan teknik pembagian (*divisive*) (Rencher, 2002). Pada penelitian ini teknik yang digunakan ialah teknik penggabungan (*agglomerative*). Teknik tersebut diperoleh dengan menggabungkan objek atau kelompok secara bertahap, sehingga pada akhirnya didapat hanya satu kelompok saja, dimana pengelompokan n objek yang diawali dengan kondisi saat semua objek terpisah (ada n kelompok) kemudian menggabungkan objek yang mirip dalam kelompok yang sama, lalu dilanjutkan menggabungkan kelompok yang mirip sampai semua objek menjadi satu kelompok (Kaufman & Rousseeuw, 1990). Metode *agglomerative* dalam algoritma hierarki menggunakan metode pautan (*linkage*). Metode pautan merupakan metode penghubung yang menggolongkan ketidakmiripan antar cluster dalam proses penyatuan. Dalam metode hierarki ada beberapa ukuran jarak antar kelompok, yaitu *single linkage* (SL), *complete linkage* (CL), *average linkage* (AL), *centroid*, *median* dan *ward*.

Ukuran jarak yang digunakan dalam penelitian ini adalah jarak *dynamic time warping* (DTW), jarak DTW digunakan untuk mengitung dan menemukan jalur terbaik antara dua data *time series* (Munthe, 2019). Algoritma DTW mengukur kesamaan antara dua data *time series* dengan sangat efisien, meminimalkan efek jeda waktu dan distorsi, memungkinkan untuk mendeteksi bahkan dalam fase yang berbeda.

$$DTW(S, T) = \min_w \left[\sum_{k=1}^p \delta(w_k) \right] \quad \dots(1)$$

dengan $S = s_1, s_2, \dots, s_i, \dots, s_n$ dan $T = t_1, t_2, \dots, t_j, \dots, t_m$ adalah suatu data deret waktu. $W = w_1, w_2, \dots, w_k$ adalah jalur lengkungan yang memetakan ulang anggota S dan T sehingga jarak keduanya minimum (Adnyani & Sihombing, 2021).

Koefisien korelasi *Cophenetic* merupakan ukuran kebaikan penggunaan suatu ketidakmiripan dalam pengelompokan data deret waktu. Koefisien korelasi *Cophenetic* adalah korelasi antara jarak *Cophenetic* yang diperoleh dari diagram pohon dengan jarak objek semula yang digunakan

untuk membuat diagram pohon (Sokal & Rohlf, 1962). Formula koefisien korelasi *Cophenetic* sebagai berikut:

$$c = \frac{\sum_{i < j} (d(i, j) - \bar{d})(v(i, j) - \bar{v})}{\sqrt{[(d(i, j) - \bar{d})^2] [(v(i, j) - \bar{v})^2]}} \quad \dots(2)$$

Dengan $d(i, j)$: ketidakmiripan antara amatan ke-i dan amatan ke-j, \bar{d} : rata-rata ketidakmiripan, $v(i, j)$: jarak antara titik dendogram ke-i dan titik dendogram ke-j, dan \bar{v} : rata-rata jarak pada dendogram.

Koefisien *silhouette* adalah salah satu ukuran yang dapat digunakan untuk menentukan jumlah kelompok (k) optimal. Koefisien ini membandingkan rata-rata kemiripan setiap objek terhadap objek lainnya dalam satu kelompok $a(i)$ dengan minimum rata-rata kemiripan setiap objek terhadap objek lainnya di luar kelompok $b(i)$. Formula yang digunakan adalah sebagai berikut:

$$s_i = \frac{b(i) - a(i)}{\max\{(i), a(i)\}} \quad \dots(3)$$

dengan $a(i)$: rata-rata jarak antar anggota dalam kelompok dan $b(i)$: jarak terkecil antara anggota kelompok dengan anggota kelompok terdekat. Adapun kategori kebaikan pengelompokan berdasarkan (Kaufman & Rousseeuw, 1990) dinyatakan pada tabel berikut:

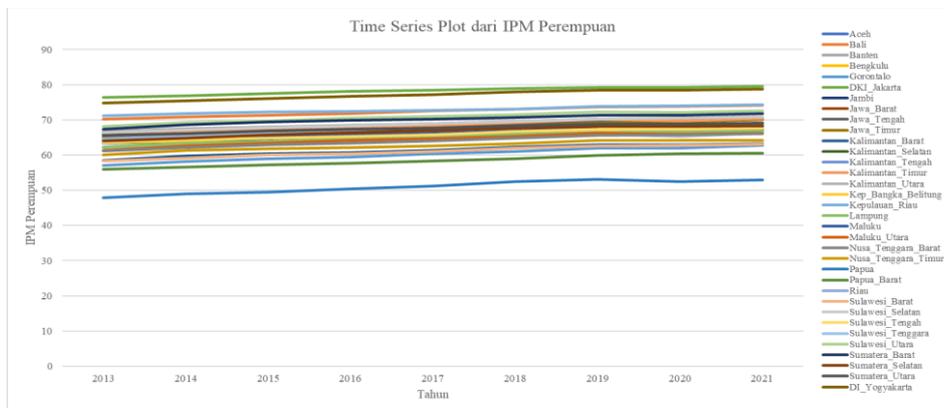
Tabel 1. Kategori Pengelompokan berdasarkan nilai Koefisien *silhouette*

| No | Koefisien <i>silhouette</i> | Kategori Pengelompokan |
|----|-----------------------------|------------------------|
| 1 | 0.71-1.00 | Kuat |
| 2 | 0.51-0.70 | Baik |
| 3 | 0.26-0.50 | Lemah |
| 4 | 0.00-0.25 | Buruk |

Nilai koefisien *silhouette* dapat digunakan untuk melihat kualitas kelompok yang dihasilkan, dan juga mengukur seberapa baik atau buruknya suatu obyek tertentu ditempatkan dalam suatu kelompok yang terbentuk (Dewi & Pramita, 2019). Semakin besar nilai koefisien *silhouette* mengindikasikan bahwa kelompok-kelompok yang terbentuk sudah terdiferensiasi dengan baik dalam artian bahwa karakteristik obyek di dalam kelompok adalah homogen dan karakteristik obyek antar kelompok heterogen.

3. HASIL DAN PEMBAHASAN

Tahapan awal dalam penelitian ini adalah melakukan eksplorasi data IPM perempuan pada 34 Provinsi di Indonesia tahun 2013-2021 dengan melihat pola *series* waktu dan ringkasan statistik pada masing-masing daerah. Hasil pola *series* waktu ditampilkan pada Gambar 1 dan ringkasan statistik ditampilkan pada Tabel 2.



Gambar 1. Pola IPM Perempuan Pada 34 Provinsi Di Indonesia Tahun 2013-2021.

Berdasarkan Gambar 1 terlihat bahwa IPM perempuan di 34 Provinsi mengalami peningkatan setiap tahunnya. Provinsi Papua terlihat mempunyai nilai IPM perempuan yang lebih rendah dibandingkan dengan daerah yang lainnya. Sedangkan daerah DKI Jakarta mempunyai nilai IPM yang paling tinggi dibandingkan dengan daerah lainnya.

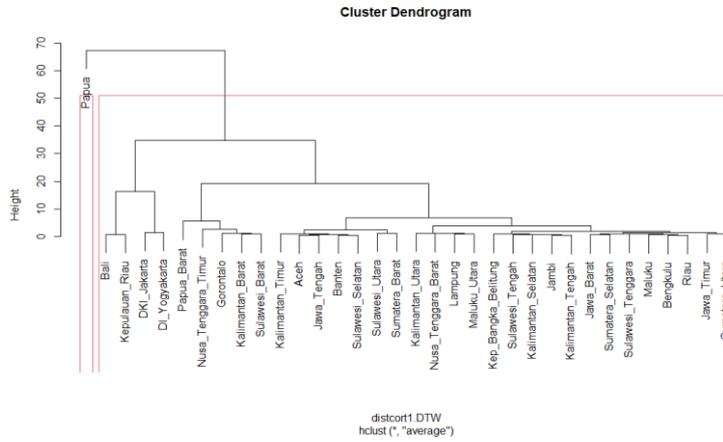
Tabel 2. Ringkasan Statistik IPM Perempuan Pada 34 Provinsi di Indonesia

| No | Provinsi | Mean | StDev | Minimum | Q1 | Median | Q3 | Maximum |
|----|---------------------|--------|-------|---------|--------|--------|--------|---------|
| 1 | Aceh | 68.37 | 1.514 | 65.68 | 67.215 | 68.41 | 69.845 | 70.18 |
| 2 | Bali | 72.393 | 1.431 | 70.2 | 71.04 | 72.62 | 73.77 | 74.17 |
| 3 | Banten | 68.919 | 1.369 | 66.65 | 67.73 | 68.99 | 70.26 | 70.56 |
| 4 | Bengkulu | 67.15 | 1.384 | 64.93 | 65.94 | 67.22 | 68.475 | 68.89 |
| 5 | Gorontalo | 60.21 | 1.923 | 57.08 | 58.56 | 60.4 | 61.965 | 62.77 |
| 6 | DKI_Jakarta | 78.228 | 1.113 | 76.35 | 77.165 | 78.48 | 79.165 | 79.54 |
| 7 | Jambi | 65.783 | 1.347 | 63.6 | 64.595 | 65.78 | 67.075 | 67.46 |
| 8 | Jawa_Barat | 66.689 | 1.447 | 64.37 | 65.37 | 66.78 | 68.055 | 68.5 |
| 9 | Jawa_Tengah | 68.497 | 1.404 | 66.1 | 67.375 | 68.48 | 69.79 | 70.36 |
| 10 | Jawa_Timur | 67.769 | 1.6 | 65.07 | 66.46 | 67.86 | 69.225 | 69.96 |
| 11 | Kalimantan_Barat | 61.357 | 1.627 | 58.55 | 60.07 | 61.38 | 62.945 | 63.22 |
| 12 | Kalimantan_Selatan | 65.611 | 1.401 | 63.39 | 64.285 | 65.76 | 66.91 | 67.35 |
| 13 | Kalimantan_Tengah | 65.658 | 1.285 | 63.43 | 64.59 | 65.74 | 66.91 | 67.06 |
| 14 | Kalimantan_Timur | 68.52 | 1.43 | 66.39 | 67.175 | 68.52 | 69.915 | 70.36 |
| 15 | Kalimantan_Utara | 64.347 | 1.431 | 62.08 | 63.135 | 64.14 | 65.66 | 66.2 |
| 16 | Kep_Bangka_Belitung | 65.907 | 1.485 | 63.58 | 64.495 | 66.02 | 67.29 | 67.69 |
| 17 | Kepulauan_Riau | 72.819 | 1.05 | 71.11 | 72.015 | 72.74 | 73.865 | 74.25 |
| 18 | Lampung | 65.03 | 1.611 | 62.21 | 63.64 | 65.24 | 66.53 | 66.81 |
| 19 | Maluku | 66.537 | 1.349 | 64.43 | 65.34 | 66.61 | 67.87 | 68.17 |
| 20 | Maluku_Utara | 64.429 | 1.727 | 61.4 | 62.96 | 64.59 | 66.1 | 66.31 |
| 21 | Nusa_Tenggara_Barat | 63.987 | 1.739 | 61.16 | 62.505 | 64.09 | 65.68 | 66.15 |
| 22 | Nusa_Tenggara_Timur | 62.659 | 1.42 | 60.12 | 61.635 | 62.58 | 64.145 | 64.22 |
| 23 | Papua | 51.009 | 1.911 | 47.88 | 49.25 | 51.24 | 52.74 | 53.14 |
| 24 | Papua_Barat | 58.38 | 1.682 | 55.93 | 56.865 | 58.26 | 60.155 | 60.52 |
| 25 | Riau | 66.991 | 1.345 | 64.68 | 65.835 | 67.17 | 68.245 | 68.46 |
| 26 | Sulawesi_Barat | 61.031 | 1.706 | 58.26 | 59.56 | 61.01 | 62.695 | 63.06 |
| 27 | Sulawesi_Selatan | 68.839 | 1.539 | 66.34 | 67.515 | 68.9 | 70.345 | 70.76 |
| 28 | Sulawesi_Tengah | 65.904 | 1.421 | 63.6 | 64.72 | 65.76 | 67.345 | 67.59 |
| 29 | Sulawesi_Tenggara | 67 | 1.603 | 64.33 | 65.67 | 67 | 68.57 | 68.98 |
| 30 | Sulawesi_Utara | 70.763 | 1.476 | 68.19 | 69.49 | 70.95 | 72.13 | 72.52 |
| 31 | Sumatera_Barat | 70.069 | 1.428 | 67.39 | 68.98 | 70.26 | 71.34 | 71.72 |
| 32 | Sumatera_Selatan | 66.709 | 1.57 | 63.95 | 65.33 | 67.08 | 68.16 | 68.34 |
| 33 | Sumatera_Utara | 67.611 | 1.252 | 65.55 | 66.52 | 67.68 | 68.815 | 69.08 |
| 34 | DI_Yogyakarta | 77.097 | 1.437 | 74.79 | 75.76 | 77.18 | 78.465 | 78.74 |

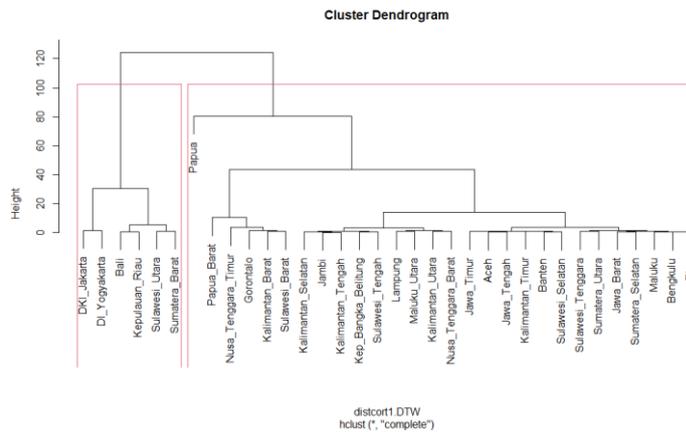
Berdasarkan Tabel 2 didapat informasi bahwa daerah yang memiliki nilai IPM perempuan tertinggi adalah DKI Jakarta dengan rata-rata IPM perempuannya sebesar 78.228. Sedangkan daerah dengan IPM perempuan terendah adalah Papua dengan rata-rata IPM perempuan sebesar 51.009. Adapun sebaran tertinggi ada di daerah Gorontalo dengan nilai standar deviasi sebesar 1.923. Hal ini mengindikasikan bahwa perubahan nilai IPM perempuan di daerah Gorontalo lebih berfluktuasi dibandingkan dengan daerah lainnya.

Langkah awal dalam melakukan analisis *clustering time series* adalah menghitung ukuran jarak. Jarak yang digunakan adalah DTW. Langkah selanjutnya adalah dengan melakukan *clustering*

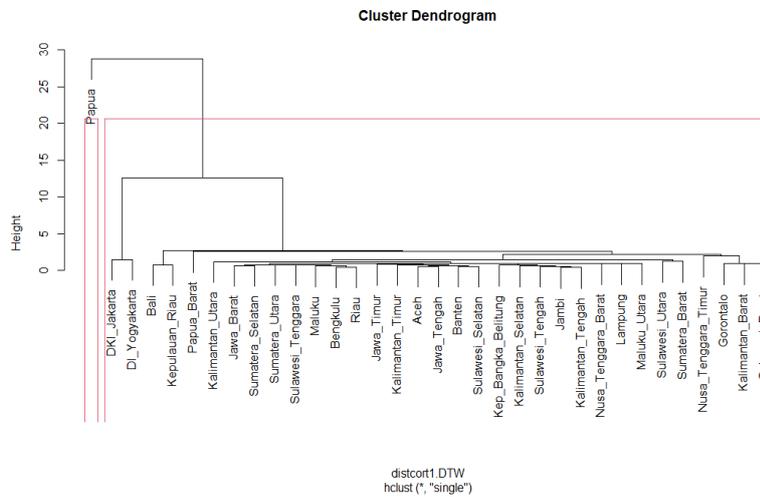
berhierarki dengan jarak DTW dan penerapan metode SL, CL, dan AL. Berdasarkan hasil analisis berikut adalah hasil penggambaran dan pemotongan dendrogram.



Gambar 2. Dendrogram *clustering* berhierarki dengan metode AL.



Gambar 3. Dendrogram *clustering* berhierarki dengan metode CL.



Gambar 4. Dendrogram *clustering* berhierarki dengan metode SL.

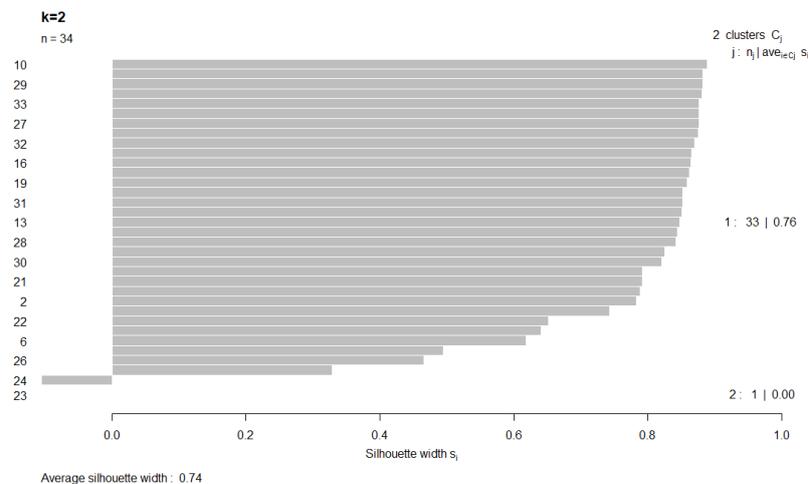
Berdasarkan Gambar 2, 3, dan 4 terlihat bahwa kelompok yang terbentuk dari dendrogram *clustering* berhierarki dengan metode AL, CL, dan SL adalah dua kelompok, dimana pada metode SL dan AL kelompok satu hanya beranggotakan daerah Papua saja dan kelompok dua beranggotakan daerah lainnya. Sedangkan pada metode CL terbentuk dua kelompok dimana kelompok satu beranggotakan daerah DKI Jakarta, DI Yogyakarta, Bali, Kepulauan Riau, Sulawesi Utara, Sumatera Barat dan kelompok dua beranggotakan daerah lainnya

Langkah berikutnya adalah menghitung nilai korelasi *Cophenetic* untuk memilih metode mana yang lebih baik, hasil perhitungan disajikan pada Tabel 3.

Tabel 3. Nilai Korelasi *Cophenetic* Ukuran Jarak DTW

| No | Linkage methods | <i>Cophenetic</i> value |
|----|-----------------|-------------------------|
| 1 | SL | 0.8123974 |
| 2 | CL | 0.6182815 |
| 3 | AL | 0.8453003 |

Berdasarkan Tabel 3, hasil nilai korelasi *Cophenetic* tertinggi diperoleh dengan menggunakan ukuran jarak DTW dan metode AL sebesar 0.8453003, sehingga untuk pengelompokan daerah berdasarkan nilai IPM perempuan yang akan diambil adalah pengelompokan menggunakan metode AL. Langkah selanjutnya adalah melakukan validasi terhadap jumlah *cluster* yang terbentuk. Hasil validasi yang dilakukan dengan menggunakan plot *silhouette* sebagai berikut:



Gambar 5. Plot *silhouette* IPM Perempuan dengan k=2.

Berdasarkan Gambar 5, diperoleh nilai rata-rata *silhouette coefficient* untuk dua kelompok sebesar 0,74 yang termasuk dalam kategori kuat berdasarkan kriteria pengukuran nilai *silhouette coefficient*. Hal ini berarti bahwa kualitas kelompok daerah berdasarkan nilai IPM perempuan sudah baik. Sehingga kelompok daerah yang terbentuk berdasarkan metode *clustering time series* dengan jarak DTW dan algoritma AL menghasilkan dua kelompok yaitu kelompok 1 daerah dengan IPM perempuan rendah yaitu daerah Papua dan kelompok 2 dengan IPM perempuan tinggi yaitu daerah lain selain Papua. Papua, yang berada di ujung timur Indonesia, memiliki banyak sumber daya alam yang potensial, tetapi belum diimbangi dengan IPM yang baik. Dilansir dari (Agapa, 2019) menurut United Nations Children's Fund (UNICEF) pada tahun 2016, 30% siswa Papua tidak menyelesaikan jenjang pendidikan SD dan SMP. Fakta ini memperlihatkan bahwa adanya kesenjangan pendidikan di daerah timur Indonesia. Hal ini sangat disayangkan dan diperlukannya perhatian pemerintah untuk menjadikan Papua sebagai salah satu daerah prioritas percepatan pembangunan di Indonesia. Berdasarkan hasil penelitian

(Yuniar & Yuniasih, 2022) menyebutkan bahwa kesetaraan gender, anggaran pemerintah, tingkat pengangguran, dan rasio gini sangat memengaruhi ketimpangan pendidikan di Indonesia. Program-program pemberdayaan perempuan di Papua sangat diperlukan guna meningkatkan keikutsertaan perempuan dalam meningkatkan pembangunan Papua di berbagai sector.

4. SIMPULAN DAN SARAN

Analisis *clustering time series* dengan menggunakan *jarak dynamic time warping* pada pengelompokan daerah di Indonesia berdasarkan IPM perempuan menghasilkan dua kelompok yaitu kelompok 1 (daerah dengan IPM perempuan rendah) yaitu daerah Papua dan kelompok 2 (daerah dengan IPM perempuan tinggi) yaitu daerah lain selain Papua. Pengelompokan yang dibentuk menghasilkan nilai koefisien *Silhouette* sebesar 0,74. Nilai tersebut menandakan bahwa kelompok yang dibentuk berada dalam kategori kuat dalam artian bahwa dua kelompok tersebut mempunyai karakteristik yang jelas berbeda sehingga metode pengelompokan dengan *jarak dynamic time warping* dapat digunakan dalam pengelompokan provinsi-provinsi di Indonesia berdasarkan nilai IPM Perempuan. Hasil pengelompokan menunjukkan bahwa diperlukannya program-program pemberdayaan bagi perempuan di Papua guna meningkatkan keikutsertaannya dalam pembangunan Papua di berbagai sektor dan juga meminimalisir kesenjangan pembangunan dari sudut pandang gender.

DAFTAR PUSTAKA

- Adnyani, L. P. W. & Sihombing, P. R., (2021). 'Analisis Cluster Time Series Dalam Pengelompokan Provinsi Di Indonesia Berdasarkan Nilai PDRB'. *Jurnal Bayesian: Jurnal Ilmiah Statistika dan Ekonometrika*, 1(1), pp. 47-54.
- Agapa, Y., (2019). 'Kesetaraan Gender dan Pemberdayaan Perempuan Papua'. [Online] Available at: <https://www.quareta.com/post/kesetaraan-gender-dan-pemberdayaan-perempuan-papua> [Accessed 3 12 2022].
- Aghabozorgi, S., Shirkhorshidi, A. S. & Wah, T. Y., (2015). 'Time-series clustering – A decade review'. *Information Systems*, Volume 53, pp. 16-38.
- Ambarwati, A. N., (2019). 'Latent Class Cluster Analysis Untuk Pengelompokan Kabupaten/Kota Di Provinsi Jawa Tengah Berdasarkan Indikator Indeks Pembangunan Manusia 2017'. *Variance : Journal of Statistics and Its Applications*, 1(2), pp. 46-54.
- Arifin, S., (2018). 'Kesetaraan Gender Dan Pertumbuhan Ekonomi Di Indonesia'. *Kajian*, 23(1), pp. 27-41.
- Dewi, D. A. I. C., & Pramita, D. A. K. (2019). 'Analisis Perbandingan Metode Elbow dan *Silhouette* pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali'. *Matrix: Jurnal Manajemen Teknologi Dan Informatika*, 9(3), 102-109.
- Kaufman, L. & Rousseeuw, P., (1990). 'Finding Groups in Data: An Introduction to Cluster Analysis'. New Jersey (US): John Wiley and Sons Inc.
- Liao, T., (2005). 'Clustering of time series data-a survey'. *Pattern Recognition*, Volume 38, p. 1857 – 1874.
- Li, L. & Prakash, B. A., (2011). 'Time Series Clustering: Complex is Simpler!'. USA, Proceedings of the 28th International Conference on Machine Learning, .
- Mattjik, A. & Sumertajaya, I., (2011). 'Sidik Peubah Ganda dengan Menggunakan SAS'. 1 ed. Bogor: IPB Press.
- Munthe, A. D. (2019). 'Penerapan *clustering time series* untuk menggerombolkan provinsi di Indonesia berdasarkan nilai produksi padi'. *Jurnal Litbang Sukowati: Media Penelitian dan Pengembangan*, 2(2), 11-11.
- Paparrizos, J. & Gravano, L., (2015). 'k-Shape: Efficient and Accurate Clustering of Time Series'. Melbourne, Proceedings of the 2015 ACM SIGMOD international conference on management of data
- Prastyo, H. E. & Ilfana, F., (2022). 'Pengelompokan Kabupaten Dan Kota Di Jawa Timur Berdasarkan Indeks Pembangunan Manusia Dengan Menggunakan Metode K-Means Tahun 2020-2021'. *JIKOSTIK – Jurnal Ilmiah Komputasi dan Statistika*, 2(1), pp. 22-32.
- Purba, R. A. et al., (2022). 'Penerapan K-Means Clustering Pada Pengelompokan Kabupaten/Kota Sumatera Utara Berdasarkan Indeks Pembangunan Manusia'. *SmartAI J O U R N A L*, 1(2), pp. 72-80.
- Rencher, A., (2002). 'Methods of Multivariate Analysis 2nd ed'. New York (US): A John Wiley & Sons, Inc..

- Sebayang, R., (2022). 'IPM Perempuan Indonesia Rendah, tapi Angka Harapan Hidupnya Tinggi'. [Online] Available at: <https://www.idntimes.com/news/indonesia/rehia-indrayanti-br-sebayang/ipm-perempuan-indonesia-rendah-tapi-angka-harapan-hidupnya-tinggi?page=all> [Accessed 24 11 2022].
- Sokal, R. & Rohlf, F., (1962). 'The Comparison of Dendograms by Objective Methods'. *International Association for Plant Taxonomy (IAPT)*, 11(2), pp. 33-40.
- Wahyudi, M. Z., (2021). 'Ketimpangan Pembangunan Manusia Masih Terjadi'. [Online] Available at: <https://www.kompas.id/baca/dikbud/2021/11/17/ketimpangan-pembangunan-manusia-masih-terjadi> [Accessed 24 11 2022].
- Wang, X., Smith, K. & Hyndman, R., (2006). 'Characteristic-Based Clustering for Time Series Data'. *Data Mining and Knowledge Discovery*, Volume 13, pp. 335-364.
- Yuniar, A. & Yuniasih, A. F., (2022). 'Pengaruh Kesetaraan Gender terhadap Ketimpangan Capaian Pendidikan di Indonesia Tahun 2015-2019'. *Jurnal Ekonomi dan Pembangunan Indonesia*, Volume 22.