

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pembangunan ekonomi di Indonesia masih menghadapi tantangan berupa ketimpangan antarwilayah. Perbedaan kinerja ekonomi antar provinsi menjadi salah satu hambatan dalam mewujudkan pertumbuhan ekonomi yang tidak hanya terpusat pada satu wilayah. Salah satu indikator yang dapat digunakan untuk melihat kondisi ekonomi suatu daerah adalah Produk Domestik Regional Bruto (PDRB) (BPS, 2024). Sebagai contoh nyata, berita resmi yang dipublikasikan oleh Badan Pusat Statistik (BPS) secara konsisten hingga tahun 2025 menunjukkan bahwa lebih dari 50% kontribusi terhadap pembentukan Produk Domestik Bruto (PDB) selalu terpusat di Pulau Jawa, sementara pulau-pulau lainnya memiliki kontribusi yang jauh lebih kecil. Ketergantungan ekonomi yang berat pada Pulau Jawa menunjukkan bahwa struktur ekonomi Indonesia masih belum seimbang.

Ketimpangan regional akan terus muncul jika pertumbuhan ekonomi tidak didorong secara merata di seluruh daerah. Ketidakseimbangan ini dapat menimbulkan berbagai dampak, mulai dari kesenjangan pendapatan, rendahnya akses layanan publik, hingga lambatnya pembangunan sosial di wilayah tertinggal. Studi yang dilakukan oleh Azim et al. (2022) menunjukkan bahwa ketimpangan ekonomi antardaerah di Indonesia umumnya dipengaruhi oleh perbedaan tingkat infrastruktur dan investasi. Oleh karena itu, memahami karakteristik ekonomi setiap provinsi menjadi langkah penting untuk memastikan penyusunan kebijakan pembangunan yang lebih tepat sasaran.

Kondisi ekonomi provinsi di Indonesia yang beragam memerlukan pendekatan analisis yang dapat mengelompokkan daerah-daerah tersebut berdasarkan pola PDRB yang serupa. Pengelompokan (*clustering*) dapat memberikan gambaran yang lebih mendalam mengenai posisi ekonomi masing-masing wilayah, serta dapat mengelompokkan provinsi-provinsi dengan karakteristik pertumbuhan yang homogen. Melalui analisis *clustering* PDRB, struktur ekonomi Indonesia dapat dipetakan ke dalam beberapa kelompok provinsi yang memiliki tingkat dan pola PDRB yang relatif serupa. Hasil pengelompokan ini dapat memberikan gambaran mengenai kondisi perekonomian Indonesia berdasarkan *cluster* yang terbentuk, mengidentifikasi kelompok provinsi dengan tingkat PDRB yang tinggi ataupun rendah, serta menilai apakah ketimpangan ekonomi antarprovinsi masih tergolong jauh.

Analisis *clustering* pada data deret waktu (*time series*) memiliki kekhususan tersendiri karena pola pergerakan suatu variabel tidak selalu berada pada titik waktu yang sama. Kondisi tersebut menjadikan *Dynamic Time Warping* (DTW) sebagai ukuran jarak yang lebih fleksibel, karena mampu menyesuaikan pergeseran waktu antar deret. Studi oleh Sabah dan Guerbaz (2024) menyatakan bahwa DTW mampu menangkap pola kemiripan yang bersifat nonlinier pada data *time series* secara lebih detail.

Salah satu tantangan lain dalam *time series clustering* adalah penentuan *centroid cluster* yang mewakili semua anggota. Sejalan dengan hal tersebut, Petitjean et al. (2010) memperkenalkan pendekatan *global averaging* yaitu *Dynamic Time Warping Barycenter Averaging* (DBA) sebagai metode untuk menghitung *centroid* deret waktu berbasis *warping path* DTW. DBA

dikembangkan sebagai respons terhadap keterbatasan teknik perataan deret waktu yang umumnya menggunakan pendekatan *pairwise averaging*, yang sangat sensitif terhadap urutan penggabungan deret waktu. Hal tersebut menyebabkan *centroid* yang dihasilkan menjadi tidak stabil dan kurang representatif, terutama ketika digunakan dalam proses iteratif seperti pada algoritma *K-Means*. Hasil penelitian tersebut juga menunjukkan bahwa *centroid* yang dihasilkan DBA memberikan hasil yang lebih baik pada seluruh *dataset* yang dicoba dibandingkan dengan *centroid* hasil pendekatan *pairwise averaging*.

Gabungan antara algoritma *K-Means*, DTW, dan DBA kemudian membentuk algoritma *K-Means Dynamic Time Warping Barycenter Averaging* (K-DBA), terbukti lebih efektif dalam menghasilkan *cluster* yang stabil dan representatif pada data *time series* karena mampu menangani pergesaeran waktu melalui penggunaan DTW sebagai pengukuran jarak dan menghasilkan *centroid* yang lebih stabil dan representatif melalui penggunaan DBA. Hal ini sejalan dengan penelitian oleh Santo *et al.* (2024) yang menerapkan K-DBA untuk membentuk *cluster* pada data kasus COVID-19 di Brazil, menyimpulkan bahwa K-DBA merupakan metode yang lebih efisien dibandingkan metode *hierarchical* dan mampu menghasilkan *cluster* yang jelas bahkan pada data kompleks. Berdasarkan uraian tersebut, pada penelitian ini akan dilakukan *clustering* pada data *time series* PDRB atas dasar harga konstan (ADHK) menurut provinsi di Indonesia pada tahun 2010-2024 menggunakan algoritma K-DBA.

Evaluasi kualitas hasil *clustering* juga menjadi aspek penting untuk memastikan bahwa *cluster* yang terbentuk memiliki tingkat pemisahan yang baik. Pada penelitian ini, kualitas hasil *clustering* dievaluasi menggunakan koefisien *silhouette*, yang mengukur sejauh mana suatu objek lebih mirip dengan anggota dalam *cluster* yang sama dibandingkan dengan *cluster* lainnya. Aghabozorgi *et al.* (2015) menyebutkan bahwa *silhouette* merupakan salah satu indeks evaluasi *cluster* yang dapat digunakan pada *time series clustering*. Studi oleh Arbelaiz *et al.* (2013) menyatakan bahwa *silhouette* merupakan salah satu indeks internal yang stabil dan informatif dalam mengevaluasi struktur *cluster*.

Berdasarkan permasalahan dan pendekatan yang digunakan, tujuan penelitian ini adalah untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan kemiripan pola pertumbuhan data *time series* PDRB Atas Dasar Harga Konstan (ADHK) periode 2010–2024 menggunakan algoritma K-DBA, serta menangkap heterogenitas pola pertumbuhan ekonomi antarprovinsi secara lebih komprehensif. Selain itu, penelitian ini bertujuan untuk mengidentifikasi karakteristik ekonomi dari masing-masing *cluster* yang terbentuk,

1.2 Rumusan Masalah

Rumusan masalah yang akan dibahas pada penelitian ini berdasarkan latar belakang yang telah dijelaskan sebelumnya, yaitu sebagai berikut:

1. Bagaimana hasil *clustering* data PDRB ADHK menurut provinsi di Indonesia pada tahun 2010-2024 menggunakan algoritma *K-Means Dynamic Time Warping Barycenter Averaging*?

2. Bagaimana karakteristik *cluster* yang dihasilkan pada data PDRB ADHK menurut provinsi di Indonesia pada tahun 2010-2024 menggunakan algoritma *K-Means Dynamic Time Warping Barycenter Averaging*?

1.3 Batasan Masalah

Penelitian ini memiliki beberapa batasan masalah agar lebih spesifik dan terarah. Batasan masalah yang digunakan pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan merupakan data PDRB ADHK menurut provinsi di Indonesia pada tahun 2010-2024.
2. Data PDRB ADHK menurut provinsi di Indonesia pada tahun 2010-2024 diperoleh dari *website* BPS (bps.go.id)
3. Algoritma *time series clustering* yang digunakan pada penelitian ini adalah *K-Means Dynamic Time Warping Barycenter Averaging*.
4. Evaluasi hasil *cluster* dilakukan menggunakan koefisien *silhouette*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang ada, tujuan penelitian yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mendapatkan hasil *clustering* data PDRB ADHK menurut provinsi di Indonesia pada tahun 2010-2024 menggunakan algoritma *K-Means Dynamic Time Warping Barycenter Averaging*.
2. Mengidentifikasi dan mendeskripsikan karakteristik masing-masing *cluster* yang terbentuk dari data PDRB ADHK menurut

provinsi di Indonesia pada tahun 2010-2024 yang dihasilkan oleh algoritma *K-Means Dynamic Time Warping Barycenter Averaging*.