

BAB I PENDAHULUAN

1.1 Latar Belakang

Analisis *cluster* merupakan teknik dalam analisis multivariat yang digunakan untuk mengelompokkan objek. Analisis *cluster* bertujuan untuk mengelompokkan objek-objek pengamatan berdasarkan karakteristik yang dimiliki. Ukuran kesamaan yang digunakan adalah ukuran jarak antar objek. Dua objek yang memiliki jarak paling dekat akan bergabung menjadi satu *cluster*. Kedekatan jarak yang dimiliki menunjukkan bahwa kedua objek tersebut memiliki kesamaan karakteristik (Yulianto dan Hidayatullah, 2014).

Terdapat dua metode umum dalam algoritma *cluster*, yaitu hirarki dan non-hirarki. Pada metode hirarki ada beberapa metode, yaitu metode *single linkage*, metode *complete linkage*, metode *average linkage*, dan metode *ward*. Pada metode non-hirarki terdapat metode *K-Means*, metode *K-Harmonic Means*, dan metode *K-modes* (Supranto, 2004).

Metode hirarki adalah suatu metode pada analisis *cluster* yang membentuk tingkatan tertentu seperti pada struktur pohon karena proses *cluster* dilakukan secara bertahap dan bertingkat. Metode hirarki dibagi berdasarkan teknik pengukuran jarak antar kelompok yang digunakan untuk mengelompokkan objek. Validitas merupakan hal yang penting di dalam analisis *cluster*, validitas digunakan untuk melihat apakah hasil analisis sudah akurat sehingga dapat menggambarkan populasinya. Metode yang dapat digunakan untuk melihat validitas suatu *cluster* yaitu metode *multiscale bootstrap*. Metode tersebut merupakan metode yang dapat digunakan untuk melihat validitas dari analisis *cluster* sebelumnya. Metode ini bekerja dengan pendekatan *bootstrap resampling* untuk setiap *cluster*. *Multiscale bootstrap resampling* digunakan untuk menghitung nilai *approximately unbiased* (AU) untuk setiap *cluster* yang terbentuk (Efron dan Tibshirani, 1993).

Pada awal tahun 2020 pandemi *coronavirus diseases* (COVID-19) mulai menginfeksi berbagai negara. Indonesia merupakan salah satu negara yang

terdampak pandemi tersebut, kasus yang pertama kali muncul pada awal maret 2020 hingga jumlah kasus terus meningkat dari waktu ke waktu. Dalam penyusunan indeks pembangunan manusia (IPM) terdapat tiga dimensi, yaitu umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak.

Berdasarkan publikasi BPS tahun 2020 menyatakan bahwa capaian dimensi umur panjang dan hidup sehat tumbuh melambat. Terdapat empat faktor tertentu yang bisa menyebabkan hal tersebut yaitu faktor lingkungan, perilaku sehat, pelayanan kesehatan, dan keturunan. Kemudian ada capaian dimensi pengetahuan tumbuh melambat yang mana pada tahun 2010 hingga 2020 memang mengalami peningkatan tetapi sejak tahun 2014 cenderung mengalami perlambatan. Dan yang terakhir dimensi standar hidup layak, mengalami penurunan dikarenakan penurunan aktivitas ekonomi yang mengakibatkan penurunan penyerapan tenaga kerja dan Tingkat Pengangguran Terbuka (TPT) berdampak terhadap peningkatannya kemiskinan.

Berdasarkan uraian yang telah dipaparkan, dilakukan penelitian untuk mengelompokkan daerah mana di Kalimantan Barat yang mengalami dampak pada penurunan umur panjang dan hidup sehat, pengetahuan dan standar hidup layak berdasarkan variabel yang mempengaruhi capaian dimensi penyusun IPM tahun 2020 yang terdampak Covid-19 dengan menggunakan metode *hierarchical clustering* dan dilanjutkan dengan metode *multiscale bootstrap* untuk melihat keakuratan hasil dari pengelompokkan tersebut.

1.2 Rumusan Masalah

Berdasarkan latar belakang, maka dapat dirumuskan permasalahan yang dapat diangkat dalam penelitian ini, yaitu Kabupaten/Kota manakah di Kalimantan Barat berdasarkan variabel yang mempengaruhi capaian dimensi IPM tahun 2020 yang terdampak Covid-19 dengan menggunakan metode *hierarchical clustering* dan dilanjutkan dengan metode *multiscale bootstrap* untuk melihat keakuratan hasil pengelompokkannya.

1.3 Tujuan

Berdasarkan rumusan masalah diatas, maka yang menjadi tujuan penelitian ini adalah mengelompokkan Kabupaten/Kota di Kalimantan Barat berdasarkan variabel yang mempengaruhi capaian dimensi IPM tahun 2020 yang terdampak Covid-19 dengan menggunakan metode *hierarchical clustering* dan dilanjutkan dengan metode *multiscale bootstrap* untuk melihat keakuratan hasil pengelompokkannya.

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini yaitu pembahasan difokuskan pada analisis *hierarchical clustering multiscale bootstrap*. Dengan metode *average linkage* untuk menghasilkan jumlah *clusternya* kemudian dilanjutkan dengan *multiscale bootstrap* untuk melihat validitas dari hasil *clusternya*. Objek yang diteliti merupakan variabel yang mempengaruhi capaian dimensi IPM tahun 2020. Data yang digunakan merupakan data seluruh Kabupaten/Kota di Kalimantan Barat pada tahun 2020.

1.5 Tinjauan Pustaka

Anugraha, G. (2005) melakukan penelitian dengan menggunakan pendekatan *multiscale bootstrap resampling* untuk memberikan ketidaktentuan dalam metode pengelompokkan *hierarchical clustering*. Data yang digunakan adalah kemiskinan di Jawa Timur. Hasil yang diperoleh adalah terdapat lima *cluster* yang terbentuk dengan kemiripan satu dengan yang lain dan signifikan. *Cluster* kelima yang terdiri dari Jember, Bondowoso, Situbondo, Probolinggo, dan Pasuruan dikatakan sebagai daerah yang cukup tinggi tingkat kemiskinannya dibandingkan dengan keempat *cluster* yang lainnya.

Ramadhani, L.; Ika P.; dan Fidia D. T. A., (2018) menggunakan metode *complete linkage* dan metode *hierarchical clustering bootstrap* pada penelitiannya. Studi kasus yang digunakan yaitu kemiskinan di Kalimantan Timur tahun 2016. Hasil *cluster* yang diperoleh dengan metode *complete linkage* sebanyak empat *cluster* dimana *cluster* pertama termasuk wilayah dengan kemiskinan tertinggi terdiri dari Balikpapan, Kutai Kartanegara, Penajam Paser

Utara, dan Bontang. *Cluster* kedua termasuk wilayah dengan kemiskinan rendah yaitu Samarinda. *Cluster* ketiga termasuk wilayah dengan kemiskinan cukup rendah terdiri dari Paser, Kutai Barat, Kutai Timur, dan Berau. *Cluster* keempat termasuk wilayah dengan kemiskinan cukup tinggi yaitu Mahakan Ulu. *Cluster* yang diperoleh dengan metode *hierarchical clustering bootstrap* sebanyak empat *cluster* dengan *cluster* pertama termasuk wilayah dengan kemiskinan terendah yaitu Samarinda dan Paser. *Cluster* kedua termasuk wilayah dengan kemiskinan tertinggi, yaitu Panajam Paser Utara, Kutai Barat, Kutai Kartanegara, Bontang, dan Kutai Timur. *Cluster* ketiga termasuk wilayah dengan tingkat kemiskinan cukup tinggi yaitu Balikpapan dan Berau. *Cluster* keempat dengan wilayah kemiskinan cukup rendah yaitu Mahakan Ulu.

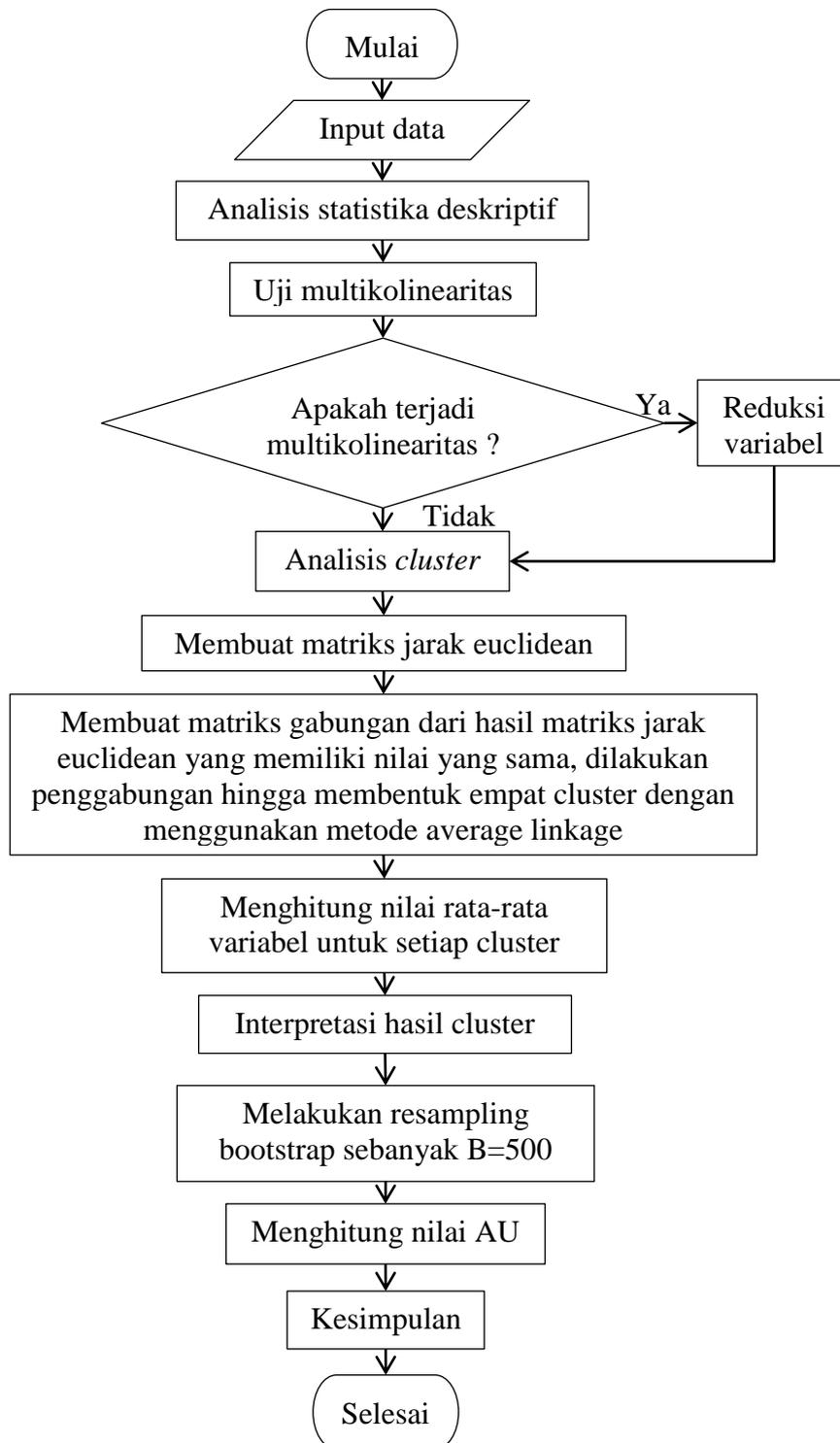
Safarina, Y.; Hayati, M. N.; Nasution, Y. N. (2019) meneliti tentang Indeks Pembangunan Manusia di Kalimantan Timur tahun 2017 dengan menggunakan penerapan metode *hierarchical clustering multiscale bootstrap*. Hasil *cluster* yang terbentuk sebanyak empat *cluster* dengan hanya tiga *cluster* yang dikatakan valid.

1.6 Metode Penelitian

Metode dan tahapan analisis yang akan digunakan pada penelitian ini dengan metode *hierarchical clustering multiscale bootstrap* sebagai berikut:

1. Menentukan rumusan masalah dalam penelitian ini
2. Analisis statistika deskriptif pada data yang digunakan untuk nantinya akan dilakukan perhitungan
3. Dilanjutkan dengan deteksi non – multikolinearitas pada data
4. Analisis *cluster* dengan bantuan *software SPSS*
5. Melakukan perhitungan *multiscale bootstrap* untuk melihat AU dengan bantuan *software R*.

Untuk lebih jelasnya dapat melihat diagram alir pada Gambar 1.1 berikut:



Gambar 1. 1 Diagram Alir *Hierarchical Clustering Multiscale Bootstrap*