

ANALISIS CLUSTER BERDASARKAN DAMPAK EKONOMI DI INDONESIA AKIBAT PANDEMI COVID-19

Ika Nur Hasanah

Program Studi Matematika, FMIPA, Universitas Negeri Surabaya

e-mail : ika.18047@mhs.unesa.ac.id

A'yunin Sofro

Program Studi Matematika, FMIPA, Universitas Negeri Surabaya

e-mail : ayuninsofro@unesa.ac.id

Abstrak

Sejak Maret 2020 Indonesia dihadapkan dengan masalah serius dengan munculnya penyakit menular yaitu Covid-19. Adanya pandemi Covid-19 menjadi masalah besar di bidang kesehatan tetapi selain itu Covid-19 juga berdampak di berbagai bidang. Kasus Covid-19 mengalami penyebaran dengan cepat, karena hal tersebut pemerintah Indonesia menetapkan berbagai kebijakan untuk menekan tingkat penyebarannya. Dengan adanya kebijakan-kebijakan tersebut berhasil menghambat penyebaran virus Covid-19, namun kebijakan tersebut juga berdampak di berbagai bidang khususnya bidang ekonomi. Dengan melihat dampak Covid-19 di bidang ekonomi, pemerintah perlu melakukan kebijakan ekonomi untuk pemulihan perekonomian di Indonesia. Untuk membuat kebijakan ekonomi, pemerintah perlu mengetahui kondisi daerah di Indonesia dan mengetahui informasi daerah-daerah di Indonesia yang sangat berdampak ekonomi akibat Covid-19. Hal ini dilakukan untuk memudahkan pemerintah dan dapat menjadi rujukan pemerintah untuk melakukan kebijakan pemulihan perekonomian di masing-masing daerah. Sehingga perlu dilakukan penelitian mengenai pengelompokan provinsi-provinsi di Indonesia berdasarkan tingkat keparahan dampak Covid-19 terhadap perekonomian menggunakan analisis *cluster*. Analisis *cluster* atau pengelompokan provinsi tersebut berdasarkan berbagai kondisi sektor ekonomi yaitu tingkat pengangguran, dan tingkat presentase penduduk miskin. Pada penelitian ini, analisis *cluster* dilakukan menggunakan metode hierarki yaitu *single linkage*, *complete linkage*, *average linkage*, dan metode *ward* dengan membandingkan menggunakan tiga ukuran jarak yaitu *Euclidean*, *Manhattan* dan *Canberra*. Hasil analisis *clustering* diperoleh metode *ward* dengan jarak *euclidean* memperoleh hasil terbaik dengan *cluster* optimal berjumlah 6, menghasilkan nilai koefisien *sillhoutte* yaitu 0,48.

Kata Kunci: Kondisi Ekonomi, Covid-19, Ukuran Jarak, *Clustering Hierarki*

Abstract

Since March 2020 Indonesia has been faced with a serious problem with the emergence of infectious disease, namely Covid-19. The Covid-19 pandemic is not only a health problem but also has an impact in various fields. Covid-19 cases are spreading rapidly, because of this the Indonesian government has set various policies to suppress the rate of spread. These policies have succeeded in inhibiting the spread of the Covid-19 virus, but these policies also have an impact in various fields, especially the economic sector. By looking at the impact of Covid-19 on the economy, the government needs to carry out economic policies for economic recovery in Indonesia. To make economic policies, the government needs to know the conditions of the regions in Indonesia and find out information about regions in Indonesia that have a very economic impact due to Covid-19. This is done to make it easier for the government and can be a reference for the government to carry out economic recovery policies in each region in Indonesia. So it is necessary to conduct research on the grouping of provinces in Indonesia based on the severity of the impact of Covid-19 on the economy using cluster analysis. The clustering of provinces is based on various economic sector conditions, namely the unemployment rate and the percentage level of the poor, which are grouped using cluster analysis. In this paper, cluster analysis was carried out using several hierarchical methods, namely *single linkage*, *complete linkage*, *average linkage*, and the *ward* method by comparing using two distance measures, namely *Euclidean*, *Manhattan* dan *Canberra*. The results of the clustering analysis obtained that the *ward* method with *Euclidean* distance obtained the best results with an optimal cluster of 6, resulting in a *sillhouette* coefficient value of 0.48.

Keywords: Economic Condition, Covid-19, Distance Measure, Hierarchical Clustering

PENDAHULUAN

Saat ini dunia dihadapkan dengan masalah serius yaitu munculnya penyakit menular yang menyebar ke hampir seluruh dunia yaitu *Coronavirus Disease 2019* atau biasa disebut Covid-19. Penyakit ini pertama teridentifikasi di Wuhan, Tiongkok (UNICEF, 2020). Dengan cepatnya penyebaran Covid-19, pada Maret 2020 Covid telah menyerang Indonesia. Hingga November 2021, kasus positif Covid-19 di Indonesia mencapai 4.255.672 orang dengan 4.103.639 orang sembuh dan 143.807 orang dinyatakan meninggal, dimana peningkatan signifikan terjadi pada bulan Juli sampai Agustus 2021 dengan kasus terbanyak pada 15 Juli 2021 yaitu sebanyak 56.757 orang terkonfirmasi positif Covid-19 (Gugus Tugas Percepatan Penanganan Covid-19, 2021).

Setelah Covid-19 menyerang Indonesia, semakin hari kasus Covid-19 di Indonesia mengalami peningkatan, hingga menyebar ke hampir seluruh provinsi di Indonesia. Karena penyebaran yang cukup cepat, pemerintah Indonesia melakukan tindakan dengan mengeluarkan berbagai kebijakan untuk menekan penyebaran Covid-19. Salah satu kebijakan yang diterapkan yaitu pemberlakuan PSBB (Pembatasan Sosial Berskala Besar) yang saat ini menjadi PPKM (Pemberlakuan Pembatasan Kegiatan Masyarakat).

Kebijakan-kebijakan yang dilakukan tersebut mampu mengurangi dan menekan tingkat penyebaran Covid-19. Namun Kebijakan-kebijakan tersebut juga membawa dampak di berbagai bidang di Indonesia khususnya di bidang ekonomi. Dengan adanya pembatasan kegiatan masyarakat selama Pandemi Covid-19 perekonomian masyarakat Indonesia menjadi terhambat dan membuat banyak karyawan mengalami pengangguran karena terkena PHK, dan dapat memicu peningkatan tingkat pengangguran dan kemiskinan di Indonesia. Hal tersebut menjadi salah satu melemahnya pertumbuhan ekonomi (Indayani & Hartono, 2020). Pada penelitian (Abd. Jalil & Kasnelly, 2020) menjelaskan mengenai pandemi Covid-19 yang berdampak pada peningkatan pengangguran bahkan diprediksi akan terus mengalami peningkatan apabila pemerintah tidak segera melakukan tindakan untuk menangani

permasalahan akibat Covid-19. Kemudian pada penelitian (Agustiana, 2020) menunjukkan mengenai dampak Covid-19. Covid-19 berdampak kepada 2,8 juta pekerja, terdiri dari 1,7 juta dan sekitar 749.499 pekerja harus kehilangan pekerjaannya karena di PHK.

Dalam menghadapi permasalahan yang serius ini, pemerintah Indonesia harus melakukan tindakan untuk memulihkan perekonomian negara agar dapat menangani krisis ekonomi setelah pandemi Covid-19. Sehingga, penelitian ini dilakukan dengan tujuan untuk mendukung pemerintah Indonesia untuk mempermudah melakukan tindakan memulihkan perekonomian. Penelitian ini bertujuan membagi daerah-daerah di Indonesia berdasarkan tingkat keparahannya berdasarkan dampak perekonomian akibat Covid-19. Variabel ekonomi yang digunakan yaitu berdasarkan akibat yang berhubungan dengan kebijakan pemerintah dalam menghadapi Covid-19.

Sebuah penelitian menyatakan bahwa pengangguran dan kemiskinan sangat mempengaruhi dan berdampak secara signifikan terhadap pertumbuhan ekonomi jangka pendek juga jangka panjang (Anderu, 2021). Beberapa penelitian juga mengatakan bahwa tingkat pengangguran maupun kemiskinan mempengaruhi pertumbuhan ekonomi (Novriansyah, 2018; Imanto dkk, 2020). Maka dari itu dua variabel ini yaitu tingkat pengangguran dan kemiskinan ini yang akan digunakan untuk melihat keadaan perekonomian di Indonesia saat kondisi Covid-19 ini.

Penelitian mengenai dampak Covid-19 dalam bidang ekonomi sebelumnya juga telah banyak dilakukan, salah satunya menganalisis tentang peran pemerintah Indonesia sebagai pembuat kebijakan untuk menangani permasalahan pandemi Covid-19 bagi UMKM (Masruroh, I. dkk, 2021). Selain itu terdapat penelitian mengkaji kebijakan pemerintah dalam menanggulangi Covid-19 dan dampaknya terhadap sosial ekonomi (Engkus dkk, 2020), dan penelitian dengan menggunakan metode deskriptif untuk menganalisis konsep kebijakan pemerintah terhadap eksternalitas ekonomi Covid-19 (Kurniawansyah dkk, 2020). Sebagian besar dari penelitian tersebut dilakukan menggunakan metode kualitatif seperti deskriptif

dan kajian teori ekonomi/kebijakan, serta belum menyentuh usaha pemulihan ekonomi.

Penelitian ini dilakukan dengan mengkaji berdasarkan analisis kuantitatif yaitu analisis *clustering*. Analisis *clustering* ini bertujuan untuk mengelompokkan daerah-daerah di Indonesia berdasarkan tingkat keparahannya berdasarkan dampak perkonomian akibat Covid-19. Analisis *clustering* adalah analisis yang bertujuan *clustering* yaitu mengelompokkan observasi ke dalam suatu *cluster* berdasarkan kesamaan (Rokach & Maimon, 2005). Analisis *clustering* secara umum memiliki dua metode. Metode pertama yaitu metode hierarki, yang terdiri *metode agglomerative* dan *devisive*. Adapun metode yang kedua yaitu metode non-hierarki. Dalam pengelompokkannya, analisis *cluster* menggunakan suatu ukuran jarak untuk medeskripsikan kedekatan antar objek/data untuk menerangkan stuktur grup sederhana dari data yang kompleks. Untuk mengukur jarak antar dua objek dalam *clustering* ada berbagai metode ukuran jarak yang dapat digunakan. Ukuran jarak yang sering digunakan dalam *clustering* yaitu ukuran jarak *Euclidean* (Johnson dan Wichern, 1982). Tetapi sebuah penelitian mengatakan bahwa Jarak *Manhattan* memiliki kinerja lebih baik dari jarak *Euclidean* (R Suwanda dkk., 2020). Selain jarak *Euclidean* dan *Manhattan* terdapat juga jarak *Canberra*, yaitu jarak yang membagi nilai mutlak selisih antar dua objek dengan jumlah dari dua objek tersebut. Dari hasil suatu eksperimen menggunakan jarak *Canberra* menunjukkan bahwa jarak *Canberra* berkinerja baik, selain dengan jarak yang sudah dikenal seperti jarak *Euclidean* (Aljohani dkk., 2017)

Penelitian sebelumnya mengenai pengelompokkan negara berdasarkan kasus Covid-19 telah dilakukan (Zarikas. V dkk, 2020) yang mana penelitian tersebut membahas *clustering* menggunakan pengelompokkan hierarki. Penelitian terkait pengelompokkan menggunakan metode jarak *Euclidean* dan *Manhattan* telah dilakukan oleh (R Suwanda dkk, 2021), dimana analisis penelitian tersebut menggunakan metode nonhierarki *K-Means*. Terkait dengan metode pengelompokkan hierarki dalam menganalisis Covid-19 di Indonesia juga telah digunakan oleh H.Fransiska dengan metode *single linkage*, *complate linkage*, dan *average linkage* dengan menggunakan jarak *Euclidean*.

Pada penelitian ini akan dilakukan pengelompokkan provinsi di Indonesia dan membandingkan hasil *cluster* menggunakan metode *hierarki single linkage*, *complate linkage*, *average linkage*, metode *ward* dengan menggunakan 3 ukura jarak yaitu jarak *Euclidean*, *Manhattan*, dan *Canberra*. Dengan penelitian ini diharapkan bisa menjadi acuan maupun rujukan untuk pemerintah Indonesia agar mempertimbangkan dalam pembuatan kebijakan ekonomi di Indonesia pada kondisi Covid-19. Hal ini bertujuan untuk memulihkan kondisi ekonomi akibat Covid-19 agar kebijakan bisa tepat pada sasaran dan sesuai dengan kondisi ekonomi yang ditargetkan.

KAJIAN TEORI

Coronavirus 2019 (Covid-19)

Coronavirus 2019 atau biasa disebut Covid-19 merupakan penyakit pernapasan yang disebabkan oleh sindrom pernapasan akut parah coronavirus 2 (SARS-CoV-2). WHO (*World Healthy Organization*) secara resmi mengumumkan bahwa Covid-19 dinyatakan sebagai pandemi. Dengan meningkatnya kasus Covid-19 di Indonesia setiap harinya, hal ini membuat pemerintah Indonesia memberlakukan kebijakan untuk meminimalisir tingkat penyebaran Covid-19. Kebijakan yang diterapkan salah satunya yaitu pemberlakuan PSBB yang kemudian berganti nama menjadi Pemberlakuan Pembatasan Kegiatan Masyarakat atau disingkat PPKM. Kebijakan tersebut terbukti mampu mengurangi dan menekan tingkat penyebaran Covid-19. Namun dengan penerapan kebijakan tersebut berdampak pada berbagai bidang, termasuk pada bidang ekonomi. Covid-19 berdampak bagi pertumbuhan ekonomi yaitu menurunnya tingkat perekonomian di Indonesia.

Kondisi pandemi Covid-19 menyebabkan permasalahan di bidang ekonomi, salah satunya berdampak pada tenaga kerja yang kehilangan pekerjaan tetap, terjadinya pemutusan hubungan kerja (PHK) sepihak sehingga menjadi pengangguran. Sehingga meningkatnya jumlah pengangguran dapat menyebabkan melemahnya pertumbuhan ekonomi. Pengangguran yaitu tenaga kerja yang sedang tidak bekerja atau bisa disebut selisih antara angkatan kerja dan tenaga kerja

sebenarnya. Sedangkan tingkat pengangguran terbuka merupakan presentase jumlah pengangguran terhadap jumlah angkatan kerja. Badan Pusat Statistik (BPS) menyatakan bahwa pada bulan Februari 2020 tingkat pengangguran terbuka di Indonesia sebesar 5,01%. Berdasarkan penelitian (Hanoatubun.S., 2020) menyatakan intervensi kuat merupakan salah satu skenario penanganan masalah Covid-19 seperti melalui PSBB yang efektif.

Cluster Analysis

Cluster analysis adalah suatu kajian atau analisis yang untuk mengklasifikasikan objek berdasarkan ciri-ciri yang hampir mirip dalam analisis statistik (Bateni dkk., 2017). Pada proses pengelompokkannya, analisis *cluster* menggunakan suatu ukuran jarak. Ukuran jarak dapat menerangkan kedekatan antar data untuk menerangkan struktur grup sederhana dari data yang kompleks.

Jarak *Euclidean* yaitu besarnya jarak suatu garis lurus yang menghubungkan antar objek. Untuk menentukan dua objek memiliki jarak terdekat, bisa dilakukan menggunakan persamaan jarak *Euclidean* yaitu mencari jarak antar objek dengan (x, y) adalah jarak antara x dan y , i adalah masing-masing data, z adalah total data, x_{ik} adalah pusat data *cluster*, dan y_{jk} adalah data pada setiap data ke- jk (Charikar dkk., 2019):

$$d(x, y) = \sqrt{\sum_{i=1}^z (x_{ik} - y_{jk})^2} \tag{1}$$

Jarak *Manhattan* adalah ukuran jarak yang digunakan untuk menghitung perbedaan multak antara koordinat sepasang objek. Jarak *Manhattan* ini dikatakan memiliki kinerja lebih baik dari jarak *Euclidean* (R Suwanda dkk., 2020). Jarak *Manhattan* didefinisikan sebagai berikut:

$$d(x, y) = \sum_{i=1}^n |x_i - y_i| \tag{2}$$

Dengan n adalah banyak data, x_j adalah pusat data *cluster*, y_j adalah data pada setiap data ke- j .

Jarak *Canberra* merupakan jumlah nilai perbedaan dibagi dengan jumlah antar dua variabel (Johnson & Wichern, 2002). Ukurn jarak ini digunakan hanya untuk variabel yang bernilai positif. Jarak *Canberra* dapat dirumuskan sebagai berikut:

$$d(x, y) = \sum_{i=1}^p \frac{|x_i - y_i|}{(x_i + y_i)} \tag{3}$$

Cluster Hierarki

Cluster Hierarki yaitu suatu metode pengelompokkan yang dari objek-objek untuk mengetahui struktur pengelompokan objek tersebut. Dalam *cluster hierarki* terdapat dua metode yaitu *agglomerative* dan *devisive*. Berikut ini merupakan prosedur metode *Agglomerative* (M. Hengki, B. Alhadi, 2015) :

1. Pengelompokkan dimulai dengan N kelompok yang setiap kelompok terdiri dari satu objek. Kemudian menghitung jarak kedekatan untuk masing-masing kelompok
2. Menghitung jarak minimum :

$$D(C_i, C_j) = \min_{1 \leq m, l \leq N, m \neq l} D(C_m, C_l) \tag{4}$$
 C_i dan C_j digabungkan untuk membentuk kelompok baru
3. Memperbarui jarak antara kelompok setelah penggabungan
4. Mengulangi langkah (2) dan (3) untuk semua elemen dalam satu kelompok.

Pada metode *agglomerative* terdapat beberapa metode keterkaitannya, salah satunya yaitu *single linkage*, *complete linkage*, *average linkage*, dan metode *ward*. Metode *single linkage* adalah metode pautan dengan menggunakan aturan jarak minimum antar *cluster*. Untuk menentukan jarak metode *single linkage* yaitu dengan melihat jarak antar dua *cluster* dan memilih jarak antar *cluster* yang terdekat (Mohbey & Thakur, 2013). Jika terdapat dua kelompok (U, V) dan W , maka rumus yang digunakan untuk menentukan jarak keduanya adalah:

$$d_{(UV)W} = \min(d_{UW}, d_{VW}) \tag{5}$$

di mana nilai d_{UW} , dan d_{VW} merupakan jarak yang paling kecil antara *cluster* U dan W serta *cluster* V dan W (Johnson & Wichern, 2007).

Complete linkage merupakan metode pengelompokkan dua objek dengan jarak terjauh. Pengelompokkan metode *complete linkage* diawali dengan menentukan objek yang mempunyai jarak paling dekat, selanjutnya menggabungkan objek tersebut dengan melihat jarak yang jauh atau maksimal (Großwendt & Röglin, 2017). Sehingga dapat dituliskan persamaan sebagai berikut :

$$d_{(UV)W} = \max(d_{UW}, d_{VW}) \tag{6}$$

dengan d_{UW} merupakan jarak yang paling jauh antara *cluster* U dan W , sedangkan d_{VW} merupakan jarak yang paling jauh antara *cluster* V dan W .

Metode *average linkage* merupakan metode *cluster* hierarki yang mengelompokkan berdasarkan rata-rata antar objek. Pengelompokkan dengan metode *average linkage* diawali dengan mendefinisikan matriks $D = (d_{ij})$ untuk menentukan objek yang dekat misal U dan V , kemudian objek tersebut digabung menjadi (UV) dan selanjutnya jarak antara (UV) dengan *cluster* lainnya W , dapat dituliskan persamaan sebagai berikut :

$$d_{(UV)W} = \frac{d_{(UV)} + d_{(VW)}}{n_{(UV)}n_W} \quad (7)$$

dengan $n_{(UV)}$ yaitu banyak anggota *cluster* (UV) dan n_W yaitu banyak anggota *cluster* W .

Metode *Ward* yaitu metode berusaha untuk meminimalkan variasi antar objek yang terdapat pada satu *cluster* (Eszergar-Kiss & Caesar, 2017). Jarak antara dua *cluster* pada metode *Ward* yaitu *sum of squares* atau jumlah kuadrat antara dua kelompok untuk semua variabel. Ukuran yang digunakan adalah *Sum of Square Error* (SSE) variabel. Berikut ini merupakan persamaan jarak yang digunakan untuk menentukan jarak dengan metode *Ward* adalah:

$$I_{ij} = \frac{1}{2} d^2(x_i, x_j) \quad (9)$$

Validasi Cluster

Validasi *cluster* merupakan tahapan untuk mengetahui tingkat ketepatan suatu *cluster*. Salah satu validasi *cluster* yang dapat digunakan yaitu koefisien *Silhouette*. Koefisien *Silhouette* merupakan suatu skala akurasi yang digunakan untuk mengetahui kualitas suatu *cluster*. Untuk menghitung koefisien *Silhouette* dapat menggunakan persamaan sebagai berikut :

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (10)$$

dengan $a(i)$ adalah jarak rata-rata antar anggota kelompok, $b(i)$ adalah nilai terkecil antara jarak rata-rata objek i dan objek pada kelompok lainnya. Nilai rata-rata *Silhouette* menunjukkan kedekatan dan kemiripan data yang dikelompokkan dalam satu *cluster* tersebut. Jika nilai rata-rata koefisien *silhouette* semakin mendekati 1 berarti berarti suatu objek sudah berada dalam *cluster* yang tepat (Putra A.K. P.,

dkk, 2015). Kriteria pengukuran nilai *Silhouette coefficient* disajikan pada tabel sebagai berikut:

Tabel 1. Kriteria *Silhouette Coefficient*

Nilai Koefisien <i>Silhouette</i>	Kriteria Cluster
0,71-1,00	<i>Strong</i>
0,51-0,70	<i>Good</i>
0,26-0,50	<i>Weak</i>
0,00-0,25	<i>Bad</i>

METODE

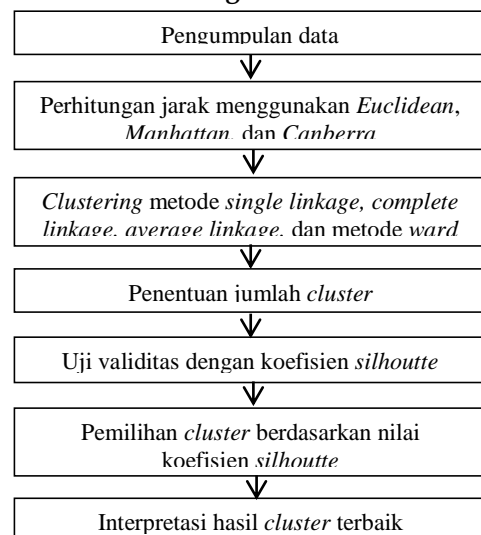
Data penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari web resmi Badan Pusat Statistika. Data yang akan dianalisis yaitu data tingkat presentase penduduk miskin tahun 2021, dan tingkat pengangguran terbuka tahun 2021 di Indonesia. Area penelitian ini adalah negara Indonesia, karena Indonesia sebagai salah satu negara yang terdampak Covid-19 di berbagai bidang khususnya di bidang ekonomi. Penelitian ini dilakukan dengan menganalisis 34 provinsi di Indonesia berdasarkan variabel ekonomi pada masa pandemi Covid-19.

Diagram Alir Penelitian

Berikut ini merupakan diagram alir yang berisi tahapan yang akan dilakukan dalam penelitian ini:

Gambar 1. Diagram Alir Penelitian



Berdasarkan Gambar 1, tahapan awal penelitian ini adalah pengumpulan data, yaitu data faktor yang mempengaruhi pertumbuhan ekonomi yaitu tingkat presentase penduduk miskin tahun 2021, dan tingkat pengangguran terbuka tahun 2021. Tahap selanjutnya melakukan perhitungan jarak dengan menggunakan jarak dengan metode *Euclidean*, *Manhattan* dan *Canberra* yang mana akan menghasilkan matriks jarak pada masing-masing metode. Setelah diperoleh matrik jarak, langkah selanjutnya yaitu melakukan *clustering hirarki* menggunakan yaitu *single linkage*, *complete linkage*, *average linkage*, dan *metode ward*. Pada proses ini dilakukan pengelompokan menggunakan tiga jarak yang berbeda yaitu pengelompokan dengan masing-masing metode *clustering* menggunakan jarak *Euclidean*, *Manhattan*, dan *Canberra*. Untuk penentuan jumlah *cluster* optimal, penelitian ini menggunakan metode *Sillhoutte*. Berdasarkan hasil jumlah *cluster* optimal, penelitian ini mengelompokkan menjadi 6 *cluster*. Hasil dari pengelompokan akan disajikan dalam bentuk dendrogram. Selanjutnya yaitu uji validitas *cluster* masing-masing metode *clustering* pada tiga jarak yang berbeda. Hal ini dilakukan untuk mengetahui kualitas pengelompokan dengan melakukan perbandingan nilai validasi untuk memperoleh *cluster* paling optimal. Pada penelitian ini validasi *cluster* menggunakan koefisien *Sillhoutte*. Nilai koefisien *silhouette* berada direntang -1 dan 1. Jika nilai *silhouette* semakin mendekati 1 maka objek sudah berada dalam *cluster* yang tepat begitu juga sebaliknya. Setelah diperoleh hasil *cluster*, langkah selanjutnya yaitu interpretasi hasil *cluster* berdasarkan dampak ekonomi di Indonesia akibat Covid-19.

HASIL DAN PEMBAHASAN

Pada bagian ini merupakan pembahasan mengenai analisis pengelompokan 34 provinsi di Indonesia berdasarkan dampak kondisi ekonomi Covid-19. Data yang digunakan yaitu data tingkat presentase penduduk miskin tahun 2021, dan tingkat pengangguran terbuka tahun 2021 di Indonesia. Berdasarkan data yang dianalisis menunjukkan

bahwa tingkat presentase penduduk miskin tahun 2021 yang tertinggi adalah Papua dengan presentase sebesar 26.86% dan presentase penduduk miskin 2021 yang terendah adalah Bali dengan presentase 4.53%. Sedangkan tingkat pengangguran terbuka tahun 2021 yang tertinggi yaitu provinsi Kepulauan Riau dengan presentase 10,12% dan Sulawesi Barat menjadi provinsi dengan tingkat pengangguran rendah yaitu 3,28%. Selain itu, diketahui pula pada Desember 2021 jumlah positif Covid-19 di Sulawesi Barat sekitar 12 ribu orang dengan presentase 0.3% dari jumlah penduduk dan termasuk kategori rendah.

Clustering berdasarkan kondisi ekonomi Covid-19

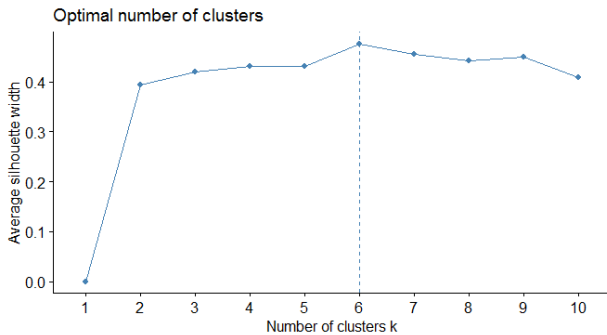
Pada penelitian ini dilakukan pengelompokan pada 34 provinsi di Indonesia yang mempunyai kesamaan karakteristik kondisi ekonomi akibat pandemi Covid-19. Tahap pertama sebelum melakukan pengelompokan yaitu dengan menentukan matriks ukuran jarak. Penelitian ini menggunakan 3 ukuran jarak yaitu *Euclidean*, *Manhattan*, dan *Canberra* yang dirumuskan seperti pada persamaan (1), (2) dan (3). Dengan bantuan R-studio didapatkan matriks ukuran jarak. Jika jarak pasangan Provinsi semakin kecil, maka semakin mirip karakteristik antar provinsi tersebut.

Langkah selanjutnya yaitu melakukan analisis metode hierarki dengan menggunakan jarak *Euclidean*, *Manhattan*, dan *Canberra* pada masing-masing metode hirarki. Pengelompokan diawali dengan menentukan objek dengan jarak paling dekat berdasarkan matriks jarak *Euclidean*, jarak paling dekat berdasarkan matriks jarak *Manhattan*, dan ukuran jarak paling dekat berdasarkan jarak *Canberra* yang telah diperoleh, kemudian menggabungkan objek tersebut dengan memperbarui jarak menggunakan perhitungan jarak masing-masing metode hierarki yang digunakan. Sehingga didapatkan *cluster* baru dan dilakukan berulang-ulang hingga mendapatkan jumlah *cluster* yang diinginkan.

Untuk menentukan jumlah *cluster*, penelitian ini menggunakan metode *silhouette*. Metode *silhouette* disini digunakan untuk menentukan jumlah *cluster* terbaik. Semakin tinggi nilai rata-rata *silhouette* maka kualitas semakin baik. Berdasarkan analisis dengan bantuan Rstudio, diperoleh bahwa jumlah *cluster* optimalnya yaitu 6

cluster. Hasil optimal cluster yaitu seperti pada Gambar 2 berikut ini:

Gambar 2. Hasil Jumlah Cluster Optimal



Hasil clustering dikelompokkan menjadi 6 cluster menggunakan metode hierarki *single linkage*, *complete linkage*, *average linkage*, dan metode *ward* yang masing-masing menggunakan jarak *Euclidean*, *Manhattan* dan *Canberra*. Dari hasil analisis, diperoleh hasil cluster dengan jarak *Euclidean* dan *Manhattan* memberikan hasil pengelompokan yang sama setiap cluster-nya. Sedangkan analisis dengan jarak *Canberra* memberikan hasil pengklasteran yang berbeda. Hasil clustering dapat dilihat dalam tabel berikut ini:

Tabel 2. Hasil Clustering Single Linkage dengan jarak Euclidean dan Manhattan

Cluster	Provinsi
1	Papua
2	NTT
3	Kep.Riau, Banten, Jawa Barat, DKI Jakarta
4	Papua Barat
5	Aceh, Maluku
6	Sumatera Barat, Riau, Jambi, Bengkulu, Lampung, , Jawa Tengah, Yogyakarta, Jawa Timur, Bali, NTB, Kep.Bangka Belitung dan 14 provinsi lainnya

Tabel 3. Hasil Clustering Complete Linkage dengan jarak Euclidean dan Manhattan

Cluster	Provinsi
1	Papua, NTT
2	Aceh, Maluku, Papua Barat
3	Kep.Riau, Banten, Jawa Barat, DKI Jakarta

4	Bengkulu, Gorontalo, NTB, Sulawesi Tengah, Lampung, Yogyakarta, Sulawesi Tenggara, Sulawesi Barat
5	Sumatera Selatan, Jawa Tengah, Jawa Timur, Sumatera Utara, Sumatera Barat, Sulawesi Utara, Sulawesi Selatan, Kalimantan Barat, Kalimantan Timur
6	Bali, Jambi, Kep.Belitung, Riau, dan 4 provinsi lainnya

Tabel 4. Hasil Clustering Average Linkage dengan jarak Euclidean dan Manhattan

Cluster	Provinsi
1	Papua
2	NTT
3	Papua Barat, Aceh, Maluku
4	Kep.Riau, Banten, Jawa Barat, DKI Jakarta
5	Bengkulu, Gorontalo, NTB, Sulawesi Tengah, Lampung, Yogyakarta, Sulawesi Tenggara, Sulawesi Barat, Sumatera Selatan, Jawa Tengah, Jawa Timur
6	Bali, Jambi, Riau, Kep.Bangka Belitung, Sulawesi Utara, Sulawesi Selatan, dan 8 provinsi lainnya

Tabel 5. Hasil Clustering metode Ward dengan jarak Euclidean dan Manhattan

Cluster	Provinsi
1	Papua, NTT
2	Aceh, Maluku, Papua Barat
3	Kep.Riau, Banten, Jawa Barat, DKI Jakarta
4	Bengkulu, Gorontalo, NTB, Sulawesi Tengah, Lampung, Yogyakarta, Sulawesi Tenggara, Sulawesi Barat
5	Sumatera Selatan, Jawa Tengah, Jawa Timur, Sumatera Utara, Sumatera Barat, Sulawesi Utara, Sulawesi Selatan, Kalimantan Barat, Kalimantan Timur
6	Bali, Jambi, Kep.Belitung, Riau, dan 4 provinsi lainnya

Tabel 6. Hasil Clustering Single Linkage dengan jarak Canberra

Cluster	Provinsi
1	Aceh, Maluku, Papua Barat

2	Papua , NTT, Bengkulu, Gorontalo, NTB, Sulawesi Tengah, Lampung, Yogyakarta, Sulawesi Tenggara, Sulawesi Barat, Sumatera Selatan, Jawa Timur
3	Jawa Tengah
4	Sumatera Barat, Banten, Kep.Riau, Jawa Barat, DKI Jakarta, Sulawesi Utara Kalimantan Timur,
5	Sumatera Utara, Sulawesi Selatan, Kalimantan Barat
6	Bali, Jambi, Kep.Bangka Belitung, Riau dan 4 provinsi lainnya

Tabel 7. Hasil Clustering Average Linkage dengan jarak Canberra

Cluster	Provinsi
1	Aceh, Maluku, Papua Barat
2	Papua, Bengkulu, Gorontalo, NTB, Sulawesi Tengah, Lampung, Yogyakarta, NTT, Sulawesi Tenggara, Sulawesi Barat
3	Jawa Tengah
4	Sumatera Selatan, Jawa Timur
5	Sumatera Barat, Banten, Kep.Riau, Jawa Barat, DKI Jakarta, Sulawesi Utara, Kalimantan Timur, Sumatera Utara, Sulawesi Selatan, Kalimantan Barat
6	Bali, Jambi, Kep.Bangka Belitung, Riau dan 4 provinsi lainnya Utara

Tabel 8. Hasil Clustering Complete Linkage dan metode Ward dengan jarak Canberra

Cluster	Provinsi
1	Aceh, Maluku, Papua Barat, Jawa Tengah
2	Papua, Bengkulu, Gorontalo, NTB, Sulawesi Tengah, Lampung, Yogyakarta, NTT, Sulawesi Tenggara, Sulawesi Barat
3	Sumatera Selatan, Jawa Timur
4	Sumatera Barat, Banten, Kep.Riau, Jawa Barat, DKI Jakarta, Sulawesi Utara, Kalimantan Timur
5	Sumatera Utara, Sulawesi Selatan, Kalimantan Barat
6	Bali, Jambi, Kep.Bangka Belitung, Riau dan 4 provinsi lainnya

Berdasarkan Tabel 2 sampai dengan Tabel 8 dapat dilihat bahwa *cluster* terbagi menjadi 6. Berdasarkan karakteristik data kondisi ekonomi di Indonesia akibat Pandemi (data tingkat presentase penduduk miskin dan tingkat pengangguran terbuka tahun 2021), dapat disimpulkan bahwa *cluster* pertama *cluster* 1 adalah provinsi dengan kondisi ekonomi yang sangat buruk, *cluster* 2 merupakan provinsi dengan kondisi ekonomi yang buruk, *cluster* 3 provinsi dengan kondisi ekonomi cukup buruk, *cluster* 4 provinsi dengan kondisi ekonomi sedang, *cluster* 5 yaitu provinsi dengan kondisi ekonomi cukup baik, dan *cluster* 6 merupakan provinsi dengan kondisi ekonomi baik.

Validasi Cluster

Tahap selanjutnya setelah dilakukan *clustering* yaitu validasi *cluster*. Penelitian ini menggunakan *koefisien silhouette*. Hal ini dilakukan untuk menentukan *cluster* optimal dengan melakukan perbandingan nilai validasi untuk

Metode	Silhouette Coefficient		
	Jarak Euclidean	Jarak Manhattan	Jarak Canberra
Single Linkage	0,3	0,27	0,45
Complete Linkage	0,47	0,44	0,46
Average Linkage	0,43	0,42	0,42
Ward Method	0,48	0,46	0,46

memperoleh *cluster* terbaik. Hasil Nilai validasi koefisien *silhouette* dapat dilihat pada Tabel 9 sebagai berikut :

Tabel 9. Nilai Validasi Clustering

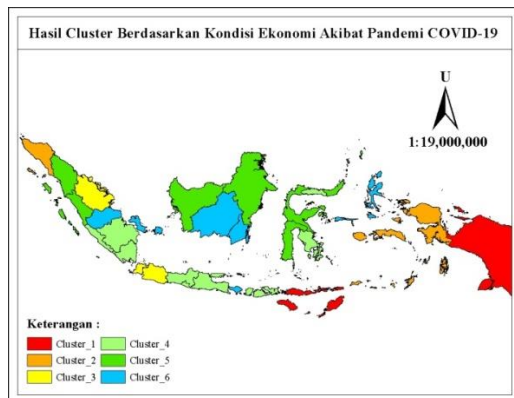
Berdasarkan nilai koefisien *silhouette* diatas, dapat diketahui bahwa pengelompokan yang paling optimal dalam mengelompokkan 34 provinsi di Indonesia yaitu pengelompokan menggunakan metode *Ward* dengan jarak *Euclidean*, nilai koefisien *silhouette*nya yaitu 0,48.

Interpretasi Hasil Cluster

Pada proses *clustering* 34 provinsi di Indonesia berdasarkan dampak ekonomi di Indonesia akibat Covid-19 menggunakan beberapa metode hirarki

dengan jarak *euclidean*, *manhattan*, dan *canberra* diperoleh hasil *cluster* yang optimal yaitu metode *ward* menggunakan jarak *euclidean* yang dibagi menjadi 6 *cluster*. Berikut ini merupakan hasil *cluster* disajikan dalam bentuk gambar peta sebagai berikut:

Gambar 3. Hasil Clustering Provinsi di Indonesia berdasarkan Kondisi Ekonomi Akibat Covid-19



Berdasarkan Gambar 3, *cluster* pertama berwarna merah terdiri dari provinsi Papua dan NTT. *Cluster* kedua berwarna orange terdiri atas provinsi Aceh, Maluku, dan Papua Barat. *Cluster* ketiga berwarna kuning terdiri atas 4 provinsi yaitu Kep.Riau, Banten, Jawa Barat, DKI Jakarta. *Cluster* keempat berwarna hijau muda terdiri atas 8 provinsi yaitu Bengkulu, Gorontalo, NTB, Sulawesi Tengah, Lampung, Yogyakarta, Sulawesi Tenggara, Sulawesi Barat. *Cluster* kelima berwarna hijau tua terdiri dari 9 provinsi yaitu Sumatera Selatan, Jawa Tengah, Jawa Timur, Sumatera Utara, Sumatera Barat, Sulawesi Utara, Sulawesi Selatan, Kalimantan Barat, Kalimantan Timur. *Cluster* enam berwarna biru terdiri atas 8 provinsi yaitu Bali, Jambi, Kep.Belitung, Riau, dan 4 provinsi lainnya

Analisis *cluster* dilakukan untuk pengklasteran provinsi-provinsi di Indonesia menurut kondisi ekonomi akibat pandemi Covid-19 yaitu dengan menggunakan data tingkat pengangguran terbuka dan tingkat presentase penduduk miskin tahun 2021. Dengan hasil analisis tersebut diharapkan dapat dijadikan pertimbangan pemerintah untuk membuat kebijakan di masing-masing provinsi, untuk memperhatikan kondisi ekonomi saat pandemi di masing-masing daerah, sehingga dapat memperbaiki keadaan ekonomi dan meminimalisir tingkat pengangguran dan tingkat kemiskinan di Indonesia.

PENUTUP

SIMPULAN

Berdasarkan hasil analisis *clustering* menggunakan beberapa metode hierarki dengan jarak *euclidean*, *manhattan*, dan *canberra* dalam pengelompokkan berdasarkan kondisi ekonomi dampak pandemi Covid-19 diperoleh metode *Ward* dengan jarak *euclidean* memperoleh hasil terbaik dibandingkan metode yang lain, dengan *cluster* optimal berjumlah 6 menghasilkan nilai koefisien silhoutte sebesar 0.48 atau dapat dikatakan kategori *Weak Clasification*.

Hasil *clustering* 34 provinsi di Indonesia berdasarkan kondisi ekonomi dampak pandemi Covid-19 menunjukkan bahwa pada *cluster* pertama terdiri dari provinsi Papua dan NTT. *Cluster* kedua terdiri atas provinsi Aceh, Maluku, dan Papua Barat. *Cluster* ketiga terdiri atas provinsi Kep.Riau, Banten, Jawa Barat, DKI Jakarta. *Cluster* keempat terdiri atas provinsi Bengkulu, Gorontalo, NTB, Sulawesi Tengah, Lampung, Yogyakarta, Sulawesi Tenggara, Sulawesi Barat. *Cluster* kelima terdiri dari provinsi Sumatera Selatan, Jawa Tengah, Jawa Timur, Sumatera Utara, Sumatera Barat, Sulawesi Utara, Sulawesi Selatan, Kalimantan Barat, Kalimantan Timur. *Cluster* enam terdiri atas 8 provinsi lainnya.

SARAN

Dalam penelitian selanjutnya, analisis *clustering* dapat dilakukan pengembangan dengan menggunakan metode *clustering* lain selain hierarki atau dapat menggunakan ukuran jarak lain seperti DTW, dan ACF dan lain sebagainya. Penggunaan metode *clustering* dan metode jarak yang berbeda ini dilakukan untuk mengetahui metode terbaik yang menghasilkan *cluster* optimal.

DAFTAR PUSTAKA

- Abd. Jalil, F., & Kasnelly, S. (2020). Meningkatnya Angka Pengangguran di Tengah Pandemi (Covid-19). *Al - Mizan : Jurnal Ekonomi Syariah*, 3(1). 45-60.
- Agustiana, L E. (2020). "Pengaruh Wabah Covid-19 Terhadap Tingkat Pengangguran Terbuka Pada Sektor Terdampak di Indonesia." *Jurnal Ilmu Manajemen Terapan*, 1(6), 546-56
- Aljohani, A., Lai, D., Bell, P., & Edirisinghe, E. (2017). A Comparison of Distance Metrics in Semi-supervised Hierarchical Clustering Methods. *Intelligent Computing Methodologies*, 719-731. doi: 10.1007/978-3-319-63315-2_63
- Anderu, K. S. (2021). An empirical nexus between poverty and unemployment on economic

- growth. *Jurnal Perspektif Pembiayaan Dan Pembangunan Daerah*, 9(1), 85 - 94. <https://doi.org/10.22437/ppd.v9i1.12005>
- Badan Pusat Statistik diakses dari www.bps.go.id/dynamic/table/2020/02/19/1774/tingkat-pengangguran-terbuka-tptmenurut-provinsi-1986---2019.html tanggal 2 Juni 2020
- Bateni, M., Behnezhad, S., Derakhshan, M., Hajiaghayi, M., Kiveris, R., Lattanzi, S., & Mirrokni, V. (2017). Affinity clustering: Hierarchical clustering at scale. *Advances in Neural Information Processing Systems*, 6864-6874.
- Bateni M, Behnezhad S, Derlaoakhshan M, Hajiaghayi MT, Kiveris R, Lattanzi S, et al. Affinity Clustering: Hierarchical Clustering at Scale. *Adv Neural Inf Process Syst*. :6864-74.
- Charikar, M., Chatziafratis, V., Niazadeh, R., & Yaroslavtsev, G. (2019, April). Hierarchical clustering for euclidean data. In *The 22nd International Conference on Artificial Intelligence and Statistics* (pp. 2721-2730). PMLR.
- Engkus, E., Suparman, N., Tri Sakti, F., & Saeful Anwar, H. (2020). Covid-19: Kebijakan mitigasi penyebaran dan dampak sosial ekonomi di Indonesia. LP2M. Retrieved from <http://digilib.uinsgd.ac.id/30820/>
- Eszergár-Kiss, D., & Caesar, B. (2017). Definition of user groups applying Ward's method. *Transportation Research Procedia*, 22, 25-34.
- Grosswendt, A., & Roeglin, H. (2017). Improved analysis of complete-linkage clustering. *Algorithmica*, 78(4), 1131-1150.
- Gugus Tugas Percepatan Penanganan COVID-19. (2021). Retrieved from <https://covid19.go.id/>
- Hanoatubun, S. (2020). Dampak Covid-19 terhadap Prekonomian Indonesia. *EduPsyCouns: Journal of Education, Psychology and Counseling*, 2(1), 146-153.
- Imanto, R., Panorama, M., & Sumantri, R. (2020). Pengaruh Pengangguran dan Kemiskinan terhadap Pertumbuhan Ekonomi di Provinsi Sumatera Selatan. *AllInfaq: Jurnal Ekonomi Islam*, 11(2), 118-139. <https://doi.org/10.32507/ajei.v11i2.636>
- Indayani, S., & Hartono, B. (2020). Analisis Pengangguran Dan Pertumbuhan Ekonomi Sebagai Akibat Pandemi Covid-19. *Jurnal Perspektif18*, (2), 201-8.
- Johnson, R.A and D.W. Wichern. (2002). *Applied Multivariate Statistical Analysis*. 5th edition. Prentice Hall. New Jersey.
- Johnson RA, Wichern DW. (2007). *Applied Multivariate Statistical Analysis*. Pearson Education, Inc.
- Johnson, R. A., & Wichern, D. W. (2014). *Applied multivariate statistical analysis* (Vol. 6). London, UK.: Pearson.
- Kurniawansyah, H., Amrullah, A., Salahuddin, M., Muslim, M., & Nurhidayati, S. (2020). Konsep Kebijakan Strategis dalam Menangani Eksternalitas Ekonomi dari Covid - 19 pada Masyarakat Rentan di Indonesia. *Indonesian Journal of Social Sciences and Humanities*, 1(2): 130-139.
- M. Hengki, B. Alhadi, L. D. (2015). Application of Hierarchical Clustering Ordered Partitioning and Collapsing Hybrid in Ebola Virus Phylogenetic Analysis. ICACSI 2015 - 2015 International Conference on Advanced Computer Science and Information Systems, Proceedings.
- Muhammad Ilham Mubarak. (2021). Retrieved 9 December 2021, from <https://muhammadilhammubarak.wordpress.com/author/mimubarakmim/>
- Mohbey, K. K., & Thakur, G. S. (2013). An experimental survey on single linkage clustering. *International Journal of Computer Applications*, 76(17), 6-11.
- Masruroh, I., Andrian, R., & Arifah, F. (2021). Peran pemerintah dalam mengatasi dampak pandemi covid-19 bagi UMKM di Indonesia. *Journal of Innovation Research and Knowledge*, 1(1), 41-48.
- Novriansyah, M. A. (2018). Pengaruh Pengangguran dan Kemiskinan Terhadap Pertumbuhan Ekonomi di Provinsi Gorontalo. *Gorontalo Development Review*, 1(1): 59-73. <https://doi.org/10.32662/golder.v1i1.115>
- Putra, A. K. P., Purwanto, Y., & Novianty, A. (2015). Analisis Sistem Deteksi Anomali Trafik Menggunakan Algoritma CURE (Clustering Using Representatives) dengan Koefisien Silhouette dalam Validasi Clustering. *eProceedings of Engineering*, 2(2).
- Rahmawati, L., Analisis Kelompok Dengan Menggunakan Metode Hierarki Untuk Pengelompokan Kabupaten/Kota Di Jawa Timur Berdasar Indikator Kesehatan, *Jurnal Matematika Vol.1 No.2 Universitas Negeri Malang*, (2012)
- Rokach, L., & Maimon, O. (2005). Clustering methods. In *Data mining and knowledge discovery handbook* (pp. 321-352). Springer, Boston, MA.
- UNICEF, Tanya-jawab seputar coronavirus (COVID-19). (2020). Retrieved from <https://www.unicef.org/indonesia/id/corona>

virus/tanya-jawab-seputar-
coronavirus?gclid=EAIaIQobChMI0KLW_y59
AIVHJNmAh15VAAfEAAYASAAEgKQkPD_
BwE#apaitunovelcoronavirus