PENGELOMPOKAN PROVINSI DI INDONESIA BERDASARKAN RESIKO COVID-19 DAN KETAHANAN PANGAN

(Clustering of Province in Indonesia Based on Risk of COVID-19 and Food Security)

Hady Suryono¹, Dwi Wahyu Triscowati²

¹Badan Pusat Statistik RI

²Badan Pusat Statistik Kabupaten Banyuwangi

E-mail: hadys@bps.go.id

ABSTRAK

Abstrak-Penyakit Coronavirus (COVID-19) adalah penyakit menular yang disebabkan virus korona serta dapat menginfeksi manusia dan menyerang sistem pernapasan. Virus ini menyebabkan infeksi pernapasan ringan maupun berat, seperti pneumonia. Persebaran kasus probable yang terjadi meningkat sangat tajam pada kurun waktu 1-2 bulan pada awal pandemi di Indonesia. Pada waktu tersebut fatality rate akibat virus COVID-19 di Indonesia mencapai 8,8%, melampaui dari negara-negara yang sudah lebih dulu terdampak virus COVID-19 lainnya seperti China (4,07%). Melihat jumlah Penduduk Indonesia yang merupakan terbesar ke-4 di dunia maka pandemi COVID-19 ini merupakan suatu ancaman yang besar. Secara luas, COVID-19 memiliki dampak terhadap faktor sosial dan ekonomi. Adanya penurunan aktivitas ekonomi terutama sektor transportasi dan perdagangan akan berpotensi terhadap ketahanan pangan seluruh wilayah di Indonesia. Dampak tersebut juga akan semakin inklusif jika melihat persebaran kasus COVID-19 yang semakin merata dan cepat di seluruh provinsi di Indonesia. Pengelompokan provinsi berdasarkan wilayah dan kesamaan karakteristik resiko COVID-19 dan ketahanan pangan diperlukan untuk mendapatkan gambaran kelompok provinsi yang perlu mendapatkan kewaspadaan yang lebih banyak. Pada penelitian ini metode K-Means digunakan untuk melakukan pengelompokan karena K-Means merupakan salah satu metode non hirarki clustering yang digunakan untuk mengelompokkan data kedalam beberapa kelompok dengan sistem partisi berdasarkan jarak antar variabel. Dari hasil analisis dipilih empat kelas untuk pengelompokkan. Dari hasil pengelompokkan, beberapa provinsi yang perlu mendapatkan perhatian lebih dipulau Jabalnusra adalah DKI Jakarta, Jawa Timur, Bali, dan Nusa Tenggara Timur, pulau sumatera adalah Riau, Kep. Riau, dan Kep. Babel, Sulawesi adalah Gorontalo dan Sulawesi Tenggara, serta daerah timur, yaitu Maluku sampai Papua.

Kata Kunci—Analisis Cluster, COVID-19, Indonesia, K-Means, Ketahanan Pangan

ABSTRACT

Coronavirus disease (COVID-19) is an infectious disease caused by coronavirus which can infect humans and attack the respiratory system. The spread of probable cases that occurred increased very sharply in the 1-2 months at the start of the pandemic in Indonesia. At that time, the fatality rate due to the COVID-19 virus in Indonesia reached 8.8%, surpassing that of countries that had already been affected by other COVID-19 viruses such as China (4.07%). Seeing the population of Indonesia which is the 4th largest in the world, the COVID-19 pandemic is a big threat. Broadly speaking, COVID-19 has an impact on social and economic factors. A decline in economic activity, especially in the transportation and warehousing sectors, will have the potential for food security in all regions of Indonesia. This impact will also be more inclusive if you look at the spread of COVID-19 cases that are increasingly even and faster in all provinces in Indonesia. The grouping of provinces based on regions and similar characteristics of the risk of COVID-19 and food security is needed to get an overview of the groups of provinces that need more awareness. In this study, the K-Means method is used for grouping because K-Means is one of the non-hierarchical clustering methods used to group data into several groups with a partition system based on the distance between variables. From the results of the analysis, four classes were selected for grouping. From the results of grouping, several provinces that need more attention on the Jabalnusra islands are DKI Jakarta, East Java, Bali, and East Nusa Tenggara, the Sumatra islands are Riau, Kep. Riau, and Kep. Babel, Sulawesi islands are Gorontalo and Southeast Sulawesi, as well as the eastern regions are Maluku to Papua.

Keywords — Cluster Analysis, COVID-19, Indonesia, K-Means, Food Security

PENDAHULUAN

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) atau Virus Corona adalah virus yang dapat menginfeksi manusia dan menyerang sistem pernapasan. Pada banyak kasus, virus ini hanya menyebabkan infeksi pernapasan ringan, seperti flu. Namun, virus ini juga bisa menyebabkan infeksi pernapasan berat, seperti infeksi paru-paru (pneumonia). Virus ini dapat menyebar antara manusia melalui percikan cairan dari mulut dan hidung (droplet) dari seseorang yang dinyatakan positif terinfeksi virus COVID-19 (kasus konfirmasi) atau dari pasien dalam pengawasan (PDP) yang saat ini dikenal kembali dengan istilah kasus suspek dan dapat juga menyebar melalui partikel-partikel kecil yang melayang di udara (*airborne*). Virus ini berawal dari Kota Wuhan di China pada 20 Januari 2020 dan masuk ke Indonesia 2 Maret 2020 dimana Presiden Republik Indonesia (RI) mengumumkan dua warga negara Indonesia (WNI) terinfeksi Corona. Persebaran kasus Probable yang terjadi meningkat sangat tajam pada kurun waktu 1-2 bulan pada awal pandemi di Indonesia. Pada waktu tersebut fatality rate atau angka kematian akibat virus COVID-19 di Indonesia mencapai 8,8%. Angka ini melampaui fatality rate di negara-negara yang sudah lebih dulu terdampak virus COVID-19 lainnya seperti China (4,07%), Amerika Serikat (3,95%) dan Iran (6,24%).

Berdasarkan data BPS, penduduk Indonesia pada Tahun 2020 adalah 271 juta jiwa. Melihat besarnya fatality rate dan juga jumlah Penduduk Indonesia yang merupakan negara dengan jumlah penduduk terbesar ke-4 di dunia maka pandemi COVID-19 yang terjadi di Indonesia merupakan suatu ancaman yang besar untuk Indonesia juga untuk peradaban dunia. Secara luas, COVID-19 memiliki dampak terhadap faktor sosial dan ekonomi yang besar (Qiu et al., 2020). Dampak tersebut juga akan semakin inklusif jika melihat persebaran kasus COVID-19 yang merata dan sangat cepat di seluruh provinsi di Indonesia yang memiliki wilayah yang sangat luas. Adanya penurunan aktivitas ekonomi terutama sektor transportasi dan perdagangan yang mengalami kontraksi 29,22 persen pada triwulan 2 2020 akan berpotensi terhadap ketahanan pangan seluruh wilayah di Indonesia.

Melihat hal tersebut, diperlukan suatu analisis pengelompokan (*cluster*) objek-objek berdasarkan kesamaan karakteristik atau relatif homogen di antara objek-objek tersebut (Walpole, 1995). Analisis *cluster* dibagi menjadi dua dilihat dari apa yang dikelompokkan, yaitu (a) pengelompokan observasi dan (b) pengelompokan variabel. Secara umum dalam pemilihan metode pengelompokan terdapat dua pendekatan yaitu pengelompokan hierarki dan pengelompokan non-hierarki. *Cluster* dapat dilakukan berdasarkan wilayah dan kesamaan karakteristik di antara provinsi-provinsi di Indonesia yang menghasilkan titik-titik pusat penyebaran kasus COVID-19. Diharapkan bahwa hasil pengelompokan dapat menunjukkan ada berapa jumlah kelompok wilayah yang signifikan berbeda jika dilihat dari data persebaran virus COVID-19. Informasi mengenai hasil klastering provinsi di Indonesia dapat membantu pemerintah untuk membuat langkah-langkah pencegahan persebaran pandemi COVID-19 agar tidak lebih meluas.

Pengelompokan provinsi di Indonesia menurut wilayah dan kesamaan karakteristik di antara provinsi-provinsi di Indonesia dengan berdasarkan kejadian COVID-19 dan potensi ketahanan pangan dilakukan dengan menggunakan metode K-means. K-Means merupakan salah satu metode klastering nonhierarki yang dapat digunakan untuk membagi objek ke dalam kelompok-kelompok berdasarkan kedekatan karakteristik, sehingga objek yang mempunyai karakteristik yang sama akan dikelompokkan ke dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda akan dikelompokkan ke dalam cluster yang lain.

Pada makalah ini akan diulas seberapa banyak kelas yang perlu digunakan dalam pengelompokan Provinsi di Indonesia berdasarkan variabel resiko COVID-19 dan ketahanan pangan menggunakan K-means, serta bagaimana gambaran hasil pengelompokan. Diharapkan dari hasil pengelompokan tersebut dapat memberikan petunjuk, provinsi mana saja yang perlu mendapatkan perhatian lebih dengan adanya kejadian COVID-19 dan ancaman masalah ketahanan pangan.

METODE

Analisis dalam penelitian ini dilakukan menggunakan R versi 3.5.1. *Package* yang digunakan adalah *cluster* untuk analisis K-means, *data.table* untuk pengolahan data seperti standarisasi data, serta *factoextra* dan *rgdal* untuk visualisasi dan pemetaan.

Variabel Penelitian

Variabel yang digunakan untuk mengelompokkan provinsi adalah variabel yang mencerminkan resiko kejadian COVID-19 dan ketahanan pangan. Variabel resiko kejadian COVID-19 menggunakan data bulan Agustus 2020, didekati dengan (1) banyaknya kejadian COVID-19 dan (2) CFR. Variabel ketahanan pangan menggunakan data tahun 2019 dan 2020, didekati dengan (1) rasio produksi padi dibandingkan jumlah penduduk, (2) kemiskinan, dan (3) kemudahan akses transportasi. Kemudahan akses transportasi di asumsikan berdasarkan pusat pulau jawa kategori mudah, sumatera dan bali nusra sedang, daerah lainnya selain Maluku dan papua sulit, Maluku dan papua sangat sulit. Oleh karena setiap variabel memiliki rentang ukuran yang berbeda, maka nilai setiap variabel distandarkan terlebih dahulu agar ukuran jarak antar kelompok tidak bias oleh ukuran rentang nilai variabel.

Analisis Cluster K-Means

K-Means merupakan salah satu metode klastering yang digunakan untuk mengelompokkan data kedalam beberapa kelompok dengan sistem partisi (Faisal et al., 2020). Metode ini masuk dalam kelompok unsupervised learning dengan sistem partisi yang memiliki tujuan untuk meminimalkan *objective function* dalam proses pengelompokan sehingga dapat meminimalkan variasi dalam satu kelompok dan memaksimalkan variasi antar kelompok. Input yang masuk pada algoritma ini berupa data tanpa label kelas. Input yang diterima adalah data dan jumlah kelompok (cluster) yang diinginkan yang kemudian algoritma akan memasukan data kedalam beberapa kelompok tersebut (Wanto, 2020). Setiap Cluster memiliki titik pusat yang disebut Centroid. K-Means sangat efektif untuk data yang berukuran besar (Johnson, R.A dan Wichern, 2007). Berikut adalah tahapan-tahapan untuk melakukan optimasi menggunakan algoritma K-Means:

- Memilih jumlah kelompok (k) yang diinginkan pada dataset
- Menentukan titik pusat (*centroid*) secara acak atau random
- Menghitung jarak terdekat setiap data dengan *centroid*. Untuk menghitung jarak terdekat dengan *centroid* menggunakan *euclidean distance* (Singh, A. Rana, A. Yadav, 2013), berdasarkan persamaan (1) berikut (Anton, 1993):

$$de = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \dots\dots\dots(1)$$

dimana:

(x, y) = koordinat obyek

(s, t) = koordinat centroid

i = banyaknya obyek

- Menghitung kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang. Pusat *cluster* adalah rata-rata dari semua data dalam sebuah *cluster*. Untuk menghitung dapat menggunakan persamaan (2) berikut:

$$V_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj} \dots\dots\dots(2)$$

dimana:

V_{ij} = Centroid rata-rata pada cluster ke - i untuk variabel ke - j

N_i = Jumlah anggota cluster ke - i

i, k = Indeks dari cluster

X_{kj} = Nilai data ke - k variabel ke - j untuk cluster tersebut

- e. Menghitung setiap objek menggunakan titik pusat *cluster* baru yang merupakan tahap awal pada iterasi berikutnya. Jika anggota *cluster* tidak mengalami perpindahan *cluster* lagi, maka proses pengelompokan dinyatakan selesai.

Virus Corona (COVID-19)

COVID-19 adalah penyakit yang disebabkan oleh virus *severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2) yang juga sering disebut virus Corona yaitu kelompok virus yang menginfeksi sistem pernapasan. COVID-19 dapat menyebabkan infeksi sistem pernapasan, mulai dari gejala yang ringan seperti flu, hingga infeksi paru-paru, seperti pneumonia (Ren, L. L. Wang, Y.M. Wu, 2020). Kasus pertama penyakit ini terjadi di kota Wuhan, Cina, pada akhir Desember 2019 yang awalnya ditularkan dari hewan ke manusia (Du Toit, 2020). Setelah itu, COVID-19 menular antarmanusia dengan sangat cepat dan menyebar ke puluhan negara, termasuk Indonesia, hanya dalam beberapa bulan. Penularan bisa melalui percikan ludah (droplet) yang keluar saat penderita COVID-19 bersin/batuk dan juga bisa menular melalui aerosol. Gejala infeksi COVID-19 muncul setelah masa inkubasi sekitar 5,2 hari (Li, Q. Guan, X. P. Wu, Wang, 2020). Periode dari timbulnya gejala COVID-19 hingga kematian berkisar antara 6 hingga 41 hari dengan median 14 hari (Wang, W. Tang, J. Wei, 2020). Periode ini tergantung pada usia pasien dan status sistem imun pasien dimana periode akan lebih pendek di antara pasien yang berusia diatas 70 tahun dibandingkan dengan mereka yang berusia di bawah 70 tahun (Wang, W. Tang, J. Wei, 2020).

HASIL DAN PEMBAHASAN

Pengelompokan Kmeans merupakan metode untuk penyusunan sub kelompok pengamatan berdasarkan jarak suatu set data, sehingga seluruh provinsi di Indonesia akan dibentuk kelompok berdasarkan kedekatan nilai variabel resiko COVID-19 dan ketahanan pangan. Untuk mengurangi bias karena ukuran antar variabel berbeda, maka setiap variabel distandarkan nilainya terlebih dahulu. Berdasarkan nilai-nilai variabel tersebut dapat dilihat masing-masing jarak antara propinsi seperti pada **Gambar 1**.

Gambar 1. Jarak Nilai Variabel antar Propinsi

Jarak antar propinsi berkisar antara 0 sampai 6. Semakin jauh jarak antar propinsi ditunjukkan oleh warna grafik yang semakin orange pekat. Sebagai contoh, jarak antara DKI Jakarta dengan Jawa Tengah berada sekitar 4, sementara jarak antara DKI Jakarta dengan Jawa Timur berkisar 2. Jarak DKI Jakarta yang paling jauh dengan Papua dan Papua Barat, yaitu sekitar

6. Semakin dekat jarak antar propinsi, meningkatkan peluang propinsi tersebut berada pada satu kelompok dan semakin jauh jaraknya, akan terpisah di kelompok yang lain.

Penentuan jumlah kelompok provinsi berdasarkan variabel penelitian dapat ditentukan berdasarkan nilai optimum jumlah antar cluster misalkan dengan jarak silhouette, elbow, ataupun GAP. Berdasarkan nilai silhouette diperoleh kelas optimum sebanyak 2 kelas, berdasarkan elbow sebanyak 7 kelas, dan berdasarkan GAP sebanyak 1 kelas. Namun penentuan jumlah kelompok juga dapat ditentukan berdasarkan kebutuhan, misal 2 atau 3 kelas saja tergantung informasi yang ingin diperoleh. Oleh karenanya, dalam pembentukan provinsi dilihat pula bagaimana komposisi kelompok menggunakan kelas sebanyak 2,3,4, dan 5 kelompok. Hasil perbandingan dapat dilihat pada **Gambar 2**.

Gambar 2. Perbandingan Kelompok Provinsi berdasarkan jumlah kelas

Terlihat dari **Gambar 2**, jika kelompok dibuat dua kelas saja, pulau Jawa mengelompok di kelompok satu. Provinsi yang ikut di kelompok satu ini adalah provinsi dengan ketahanan pangan yang baik dan akses transportasi tidak terlalu sulit. Ketika kelompok dibuat menjadi 3,4, atau 5 kelas, provinsi DKI Jakarta dan Jawa Timur selalu mengelompok menjadi satu kelas. Pada kelas sebanyak tiga, nampak bahwa Papua, Sulawesi, dan Kalimantan menjadi satu kelompok. Saat kelas ditingkatkan menjadi empat kelas, Papua menjadi kelompok sendiri bergabung dengan Papua Barat, Maluku, Nusa Tenggara Timur, dan Gorontalo. Saat jumlah kelas ditingkatkan menjadi lima kelas, tidak ada perbedaan anggota kelompok yang signifikan dengan kelas sebanyak empat kelas, sehingga pemilihan kelas sebanyak empat kelas dianggap lebih tepat. Kelompok dengan empat kelas ini dapat memberikan informasi yang lebih banyak daripada saat pengelompokan dilakukan menggunakan dua kelas saja.

Nilai rata-rata pusat dari setiap kelompok dapat dilihat pada **Tabel 1**, sedangkan hasil pengelompokan dapat dilihat pada **Gambar 3**. Nilai rata-rata setiap kelompok ini mencerminkan pusat nilai variabel setiap kelompok. Misalkan pada kelompok satu, jumlah COVID-19 memusat sekitar 24026 kejadian, CFR sebesar 5.65, persentase kemiskinan sekitar 7.81 persen, akses transportasi mudah, dan rasio produksi padi/jumlah penduduk sebesar 129.

Tabel 1. Nilai Rataan Variabel Setiap Kelompok

Kelompok	Jumlah COVID-19	CFR	Persentase Kemiskinan	Tingkat Kemudahan Transportasi	Rasio Produksi Padi / Jumlah Penduduk
1	24026	5.65	7.81	1	129
2	3132	4.5	10	2.12	262
3	1298	1.54	20.3	3.4	109
4	1007	1.93	6.65	2.6	85.6

Dilihat dari **Gambar 3**, kelompok pertama terdiri dari Provinsi DKI Jakarta dan Jawa Timur saja. Kedua provinsi ini memang diketahui sama-sama memiliki kejadian COVID-19 yang jauh lebih banyak dari propinsi lain dan jumlah penduduk yang banyak. Namun dua provinsi ini memiliki akses transportasi yang masih mudah serta secara posisi geografis masih bertetangga dengan Jawa Tengah, Jawa Barat, Yogyakarta, dan Banten yang berada di kelompok 2 dengan ciri ketahanan pangan paling tinggi. Oleh karenanya kewaspadaan terhadap dua provinsi ini masih dapat ditopang oleh provinsi di sekitarnya.

a) Kelompok Provinsi Berdasarkan Dimensi Variabel

b) Kelompok Provinsi Berdasarkan Peta Wilayah

Gambar 3. Kelompok Provinsi Berdasarkan Variabel resiko COVID-19 dan Ketahanan Pangan

Papua dan Papua Barat menjadi satu kelompok di kelompok ketiga bersama dengan Gorontalo, Nusa Tenggara Timur, dan Maluku. Gorontalo merupakan salah satu kota yang maju di daerah pulau Sulawesi, namun ternyata provinsi perlu mendapatkan perhatian khusus karena ikut terpusat di kelompok tiga ini. Ciri kelompok tiga ini adalah jumlah COVID-19 sedang mendekati rendah, CFR rendah, kemiskinan tinggi, kemudahan akses transportasi sulit, dan rasio produksi padi / jumlah penduduk sedang mendekati rendah. Jadi kelompok ketiga ini dicerminkan dengan variabel resiko COVID-19 yang rendah, namun perlu perhatian banyak untuk ketahanan pangan ketika kejadian COVID-19 meningkat dan menyebabkan akses transportasi dan pergudangan terganggu

Kelompok kedua merupakan kelompok yang dapat menjadi penyangga kelompok lain, terdiri dari 18 provinsi. Kelompok kedua ini memiliki ciri kejadian COVID-19 yang sedang cenderung rendah dan ketahanan pangan tinggi. Contohnya adalah provinsi di Jawa selain DKI dan Jawa Timur, serta beberapa provinsi di Sumatera, Kalimantan, dan Sulawesi.

Hasil pengelompokan provinsi berdasarkan variabel resiko COVID-19 dan ketahanan pangan ini diharapkan dapat memberikan warning awal terhadap provinsi yang perlu diberi perhatian. Misalkan di pulau Jawa yang perlu mendapatkan perhatian adalah DKI Jakarta dan Jawa Timur, namun kedua provinsi ini dapat disangga oleh provinsi Jawa Barat, Jawa Tengah, Yogyakarta, atau Banten. Di Pulau Sulawesi yang perlu menjadi perhatian utama adalah Gorontalo dan Sulawesi Tenggara. Provinsi Bali pun yang merupakan provinsi maju juga perlu mendapatkan perhatian jika sewaktu-waktu kejadian COVID-19 meningkat sehingga transportasi dan pergudangan antar daerah terganggu.

KESIMPULAN

Dari beberapa kriteria penentuan kelas optimal, yaitu silhouette, elbow, dan gap, diperoleh jumlah kelas optimal yang berbeda-beda, yaitu 1,2, dan 7 kelas. Penentuan jumlah kelas selain dengan metode kelas optimum, juga dapat dilakukan dengan pemilihan informasi apa yang hendak diberikan berdasarkan banyaknya kelas yang terbentuk, karena semakin banyak kelas belum tentu memberikan informasi lebih baik. Oleh karenanya dicobakan pengelompokan provinsi menggunakan 2,3,4, dan 5 kelas. Dari hasil tersebut dipilih kelas sebanyak empat karena empat kelompok memberikan informasi lebih banyak daripada dua atau tiga kelompok, serta tidak ada perbedaan signifikan antara empat kelompok dengan lima kelompok.

Berdasarkan empat kelompok yang terbentuk, diperoleh informasi bahwa semua provinsi di pulau Jawa kecuali DKI Jakarta dan Jawa Timur, serta beberapa provinsi di pulau lain terkumpul dalam kelompok dua dengan ciri tingkat kejadian COVID-19 sedang cenderung rendah dan ketahanan pangan yang baik. Kelompok dua ini dapat menjadi penopang bagi ancaman ketahanan pangan di provinsi lain. Kelompok yang paling perlu mendapatkan perhatian adalah kelompok tiga, yaitu provinsi Papua, Papua Barat, Maluku, Gorontalo, dan Nusa Tenggara Timur. Kejadian COVID-19 di kelompok tiga ini cenderung rendah, namun ketahanan pangan rendah dan akses transportasi sulit. Akibatnya jika kejadian COVID-19 di provinsi lain meningkat, ketahanan pangan pada provinsi kelompok tiga ini juga akan terdampak.

DAFTAR PUSTAKA

- Anton, H. (1993). *Elementary Linear Algebra* (7th ed.). Wiley.
- Du Toit, A. (2020). Outbreak of a novel coronavirus. *Nat. Rev. Microbiol*, 18, 123. <https://doi.org/10.1038/s41579-020-0332-0>
- Faisal, M., Zamzami, E. M., & Sutarman. (2020). Comparative Analysis of Inter-Centroid K-Means Performance using Euclidean Distance, Canberra Distance and Manhattan Distance. *Journal of Physics: Conference Series*, 1566(1). <https://doi.org/10.1088/1742-6596/1566/1/012112>
- Johnson, R.A dan Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. Pearson Education, Inc.
- Li, Q. Guan, X. P. Wu, Wang, X. (2020). Early transmission dynamics in wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2001316>
- Qiu, Y., Chen, X., & Shi, W. (2020). Impacts of social and economic factors on the transmission of coronavirus disease 2019 (COVID-19) in China. *Journal of Population Economics*, 33(4), 1127–1172. <https://doi.org/10.1007/s00148-020-00778-2>
- Ren, L. L. Wang, Y.M. Wu, Z. Q. (2020). Identification of a novel coronavirus causing severe pneumonia in

- human: a descriptive study. *Chinese Med J.* <https://doi.org/10.1097/CM9.0000000000000722>
- Singh, A. Rana, A. Yadav, A. (2013). K-means with Three different Distance Metrics. *Int. J. Comput. Appl.*, 67, 13–17.
- Walpole, R. E. (1995). *Pengantar statistika*. PT Gramedia Pustaka Utama.
- Wang, W. Tang, J. Wei, F. (2020). Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. *J. Med. Virol.*, 92(4), 441–447. <https://doi.org/10.1002/jmv.25689>
- Wanto, A. (2020). *Data Mining: Algoritma dan Implementasi* (1st ed.). Yayasan Kita Menulis.