

ABSTRAK

COVID-19 merupakan penyakit yang sedang mewabah dan menggemparkan dunia. Virus ini pertama kali muncul di Kota Wuhan, China. Diidentifikasi genetiknya pada tanggal 11 Januari 2020. Penyebaran virus ini melalui *droplet* atau cipratan air dari hidung maupun mulut saat orang yang positif COVID-19 berbicara, bersin, maupun batuk dan *droplet* tersebut masuk melalui mata, hidung, juga mulut orang lain. Penyebarannya dapat secara tidak langsung yaitu melalui benda yang terdapat virus SARS-CoV-2 dan disentuh oleh orang, lalu orang tersebut menyentuh hidung, mata, maupun mulutnya. Setiap negara memiliki persentase penduduk yang terinfeksi berbeda-beda jumlahnya. Penelitian ini bertujuan untuk mengelompokan negara berdasarkan angka kasus dan kematian baru terhadap COVID-19. Pengelompokan negara menggunakan data yang diunduh pada situs resmi World Health Organization (WHO) dan menggunakan rentang waktu 14 Maret 2020 sampai 14 April 2020. Metode yang digunakan pada penelitian ini adalah *K-Means Clustering*. Data dibagi menjadi beberapa kelompok berdasarkan kemiripan data pada atribut tertentu lalu hasil akhirnya ditentukan menggunakan *Silhouette Coefficient*. Hasil pengujian pada variasi *cluster* terbaik adalah 3 *cluster* dengan hasil akhir nilai *Silhouette* pada tanggal 14 Maret 2020 sampai 13 April 2020 sebesar 0,971, tanggal 7 Juni 2020 sampai 6 Juli 2020 sebesar 0,955, dan tanggal 3 Januari 2021 sampai 2 Februari 2021 sebesar 0,919.

Kata Kunci: Wabah, COVID-19, *Clustering*, *K-Means*, Pengelompokan Negara, *Silhouette Coefficient*

ABSTRACT

COVID-19 is a disease that is spreading and shaking the world. This virus first appeared in the city of Wuhan, China. Genetically identified on January 11, 2020. The spread of this virus is through droplets or splashes of water from the nose or mouth when a person who is positive for COVID-19 talks, sneezes, or coughs and the droplets enter through the eyes, nose, and mouth of other people. The spread can be indirect, namely through objects that contain the SARS-CoV-2 virus and are touched by people, then the person touches their nose, eyes, or mouth. Each country has a different percentage of the infected population. This study aims to categorize countries based on the number of new cases and deaths against COVID-19. The grouping of countries uses data downloaded on the official website of the World Health Organization (WHO) and uses a time span of March 14, 2020 to April 14, 2020. The method used in this study is K-Means Clustering. The data is divided into several groups based on the similarity of the data on certain attributes and the final result is determined using the Silhouette Coefficient. The test results on the best cluster variation are 3 clusters with the final Silhouette value on March 14, 2020 to April 13, 2020 of 0.971, June 7, 2020 to July 6, 2020 of 0.955, and January 3, 2021 to February 2, 2021 of 0.919.

Keywords: Outbreak, COVID-19, Clustering, K-Means, Country Grouping, Silhouette Coefficient