

## BAB I

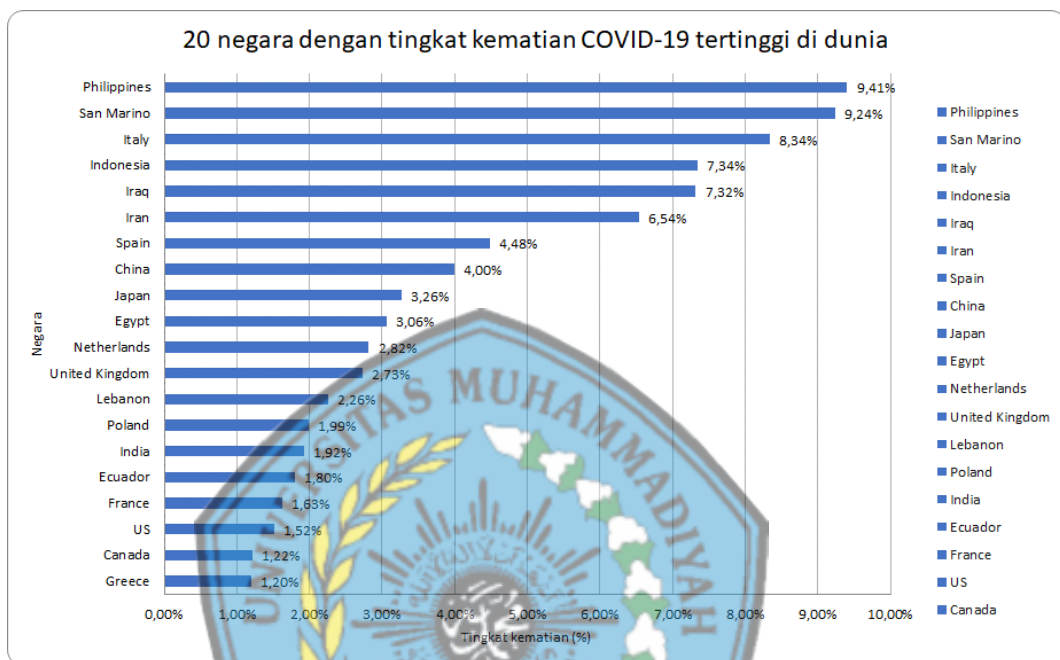
### PENDAHULUAN

#### 1.1.Latar Belakang

Indonesia sedang dilanda pandemi Covid-19, virus ini pertama kali muncul di Wilayah Wuhan-tiongkok ditemukan pada akhir Desember 2019. *Corona Virus Disease* 2019 yang disingkat dengan Covid-19 merupakan penyakit yang baru ditemukan dan disebabkan oleh jenis Corona virus baru yaitu Sars-CoV-2. Virus ini menyebar dengan cepat ke berbagai Negara hampir ke seluruh dunia sehingga *World Health Organization* (WHO) menyatakan Covid-19 sebagai *pandemic* global yang secara resmi diumumkan pada tanggal 11 Maret 2020 (WHO, 2020) dan menjadi salah satu kejadian luar biasa yang tidak pernah diperkirakan sebelumnya. Covid-19 dapat berpindah secara tidak langsung dengan menyebarkan virusnya melalui benda-benda mati yang bertahan selama dua jam sampai enam hari akibat terpapar virus dari sentuhan tangan seseorang yang terpapar. Pemerintah dan masyarakat sedang melakukan *social distancing* guna memutus rantai penyebaran Covid-19 yang terus meningkat setiap harinya. Penularan Covid-19 hanya bisa dicegah dengan kedisiplinan untuk tetap berada di rumah.

Covid-19 menyebar dengan cepat ke seluruh dunia salah satunya Indonesia. Indonesia mengumumkan kasus pertama Covid-19 pada tanggal 2 Maret 2020 dengan penemuan 2 kasus di Depok. Jumlah ini terus meningkat dengan cepat

hingga tanggal 19 Desember terdata total kasus Covid-19 terkonfirmasi sebanyak 657.948 jiwa, sembuh sebanyak 536.260 jiwa dan total kasus meninggal sebanyak 19.659 jiwa. (covid19.go.id).



Gambar 1.1. Jumlah 20 Negara Dengan Tingkat Kematian Covid-19 Tertinggi di Dunia

Sumber: (catchmeup.id/covid-19)

Berdasarkan gambar 1.1 di atas merupakan gambaran Jumlah 20 Negara dengan tingkat kematian Covid-19 tertinggi di dunia. Berdasarkan grafik tersebut bahwa Negara Philliphines menjadi urutan pertama dengan tingkat kematian tertinggi di dunia dengan jumlah kasus Covid-19 sebesar 9,41%, di urutan kedua ada San Marino dengan jumlah kasus Covid-19 sebesar 9,24%, di urutan ketiga Italy dengan jumlah kasus Covid-19 sebesar 8,34%, di urutan keempat ada Indonesia dengan jumlah kasus Covid19 sebesar 7,34%, kelima ada Iraq dengan jumlah kasus Covid sebesar 7,32% dan selanjutnya Negara-negara di dunia

lainnya berdasarkan Gambar 1.1 Jumlah 20 Negara dengan tingkat kematian Covid-19 tertinggi di dunia di atas.

*Data mining* adalah sebuah proses pencarian secara otomatis informasi yang berguna dalam tempat penyimpanan data berukuran besar. *Data mining* adalah analisa terhadap data untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut. Teknik *data mining* yang termasuk *descriptive mining* adalah *clustering*, *asosiation*, dan *sequential mining*.

*Clustering* adalah mengelompokkan data (objek) yang didasarkan hanya pada informasi yang ditemukan dalam data yang menggambarkan objek tersebut dan hubungan antar data (Tan et al., 2006). Tujuan analisis kelompok adalah agar objek-objek yang bergabung dalam sebuah kelompok merupakan objek-objek yang mirip (atau berhubungan) satu sama lain dan berbeda (atau tidak berhubungan) dengan objek dalam kelompok yang lain (Prasetyo, 2012).

Terdapat banyak metode yang bisa digunakan untuk mengelompokkan data, di antaranya ada K- Means, pengelompokkan hierarki, Fuzzy C-Means, self-organizing map, dan DBSCAN. *Density-Based Spatial Clustering Application With Noise* (DBSCAN) adalah algoritma pengelompokan yang didasarkan pada kepadatan (*density*) data. Konsep kepadatan dalam DBSCAN menghasilkan tiga macam status dari setiap data, yaitu inti (*core*), batas (*border*), dan noise. Jumlah kelompok yang terbentuk bergantung pada dua nilai yaitu Epsilon (*Eps*) yang merupakan radius (*jarak maksimal*) data anggota kelompok dari data inti, dan Minimal Points (*MinPts*) yang merupakan banyaknya minimal data kelompok

dalam radius *Eps*. DBSCAN masuk kedalam pengelompokan sekatan (*partitioning*) dan *eksklusif*, yang artinya DBSCAN merupakan metode pengelompokan yang membagi set data ke dalam sejumlah kelompok yang tidak tumpang-tindih antara satu kelompok dengan kelompok yang lain. DBSCAN dapat mendeteksi data yang mempunyai jarak yang jauh dari suatu kelompok (*noise*) yaitu data yang tidak ikut dalam kelompok mayoritas (Tan et al., 2006).

Salah satu penelitian terdahulu terkait Metode *Density-Based Spatial Clustering Application With Noise* (DBSCAN) pernah dilakukan oleh ahamsyah (2020) untuk mengelompokkan penyebaran peting di Wilayah Pasuruan pada tahun 2018. Hasil yang didapat yaitu terdapat 2 cluster dan 15 titik *noise*, dengan anggota *cluster* pertama adalah lokasi yang paling banyak terjadi petir dan sangat rapat, berjumlah 25535 titik sambaran yang menyebar hampir di seluruh Wilayah Pasuruan kecuali Prigen dan sekitarnya, anggota kedua adalah lokasi petir yang cukup rapat, namun kejadiannya tidak sebanyak petir pada *cluster* pertama, berjumlah 16 titik sambaran yang berada di sekitar Prigen. Sedangkan *noise* berada di sekitaran *cluster* kedua dan beberapa titik di sepanjang garis lintang - 7,7 derajat.

Peneliti mengenai terkait *Density-Based Spatial Clustering Application With Noise* (DBSCAN) dan K-Means pernah dilakukan oleh Isnarwaty dkk (2019) untuk mengelompokkan pendapat menjadi beberapa kategori. Hasil yang didapat yaitu nilai *silhouette coefficient* tertinggi yang diperoleh, didapatkan hasil bahwa metode DBSCAN merupakan metode terbaik dibandingkan dengan metode K-Means untuk mengelompokkan tweet yang ditujukan kepada akun media sosial

Twitter layanan ekspedisi JNE, J&T, dan Pos Indonesia. *Clustering* dengan metode terbaik menghasilkan 18 cluster untuk layanan ekspedisi JNE, 22 cluster untuk layanan ekspedisi J&T, dan 11 cluster untuk layanan ekspedisi Pos Indonesia. Kemudian ada penelitian tentang Covid-19 yang dilakukan oleh Sindi dkk (2020) menggunakan metode K-Medoids, pengklasteran terbaik dilakukan dengan 3 *cluster*. Dari 34 record diperoleh 1 record pada cluster pertama, 2 record pada *cluster* kedua, 31 record pada cluster ketiga.

Dari uraian di atas, peneliti ingin menerapkan metode *Density-Based Spatial Clustering Application With Noise* (DBSCAN) pada perkembangan kasus Covid-19 di Indonesia untuk mengetahui pola pemilihan penentuan pengelompokan penyebaran Covid-19 di masing-masing Wilayah di Indonesia

## 1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut :

1. Bagaimana cluster yang terbentuk pada daerah yang terdampak Covid-19 di masing-masing Provinsi di Indonesia menggunakan Metode DBSCAN ?
2. Bagaimana deskripsi cluster pada daerah yang terdampak Covid-19 di masing-masing Provinsi di Indonesia menggunakan Metode DBSCAN?

### 1.3. Tujuan Penelitian

Berdasarkan rumusan masalah diatas tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut :

1. Cluster yang terbentuk pada daerah yang terdampak Covid-19 di Provinsi Indonesia menggunakan Metode DBSCAN.
2. Deskripsi cluster pada daerah yang terdampak Covid-19 di masing-masing Provinsi di Indonesia menggunakan Metode DBSCAN.

### 1.4. Manfaat Penelitian

Adapun manfaat yang bisa dipetik dari penelitian ini diantaranya adalah:

#### 1. Manfaat Teoritis

Membantu perkembangan ilmu pengetahuan mengenai metode pengelompokan sehingga dapat digunakan sebagai bahan bacaan dan referensi bagi pembaca dalam melakukan peramalan terutama menggunakan metode *Density-Based Spatial Clustering Application With Noise* (DBSCAN).

#### 2. Manfaat Praktis

- a. Bagi Peneliti, dapat mengembangkan ilmu statistik yang diperoleh selama perkuliahan pada kasus yang terjadi di masyarakat. Dalam hal ini yaitu penerapan metode *Density-Based Spatial Clustering Application With Noise* (DBSCAN) pada kasus Covid-19 di Indonesia untuk mengetahui pola pemilihan penentuan

pengelompokan penyebaran Covid-19 di masing-masing Provinsi di Indonesia.

- b. Bagi pihak terkait diantaranya pemerintah dan masyarakat yaitu dapat membantu mengetahui pengelompokan penyebaran Covid-19 di Provinsi Indonesia sehingga dapat merencanakan dan mengambil keputusan yang tepat untuk melakukan berbagai aktivitas.

### 1.5. Batasan Masalah

1. Data yang digunakan dalam penelitian ini adalah data penyebaran Covid-19 di masing-masing provinsi di Indonesia meliputi kasus positif, sembuh, dan meninggal.
2. Kriteria kebaikan cluster diukur dengan menggunakan nilai *Silhouette Koefisien*.

