

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

*Covid-19* adalah penyakit menular yang disebabkan oleh sindrom pernafasan akut dari virus *SARS CoV2*. Gejala *Covid-19* bervariasi, tetapi biasanya termasuk demam, batuk, sakit kepala, kelelahan, sesak napas, dan kehilangan penciuman dan perasa. Kasus pertama yang diketahui teridentifikasi di Wuhan, China pada Desember 2019. Penyakit tersebut telah menyebar ke seluruh dunia sehingga menyebabkan pandemi yang berkelanjutan, termasuk di Indonesia. Penyebaran *Covid-19* dapat mencakup berbagai cara komunikasi, antara lain melalui droplet saat seseorang batuk, bersin, berbicara, dan bernapas. Selain itu, penyebaran *Covid-19* juga dapat terjadi melalui udara, permukaan yang terkontaminasi virus, dan kotoran manusia itu sendiri (Ramadanti and Muslih, 2021).

Jumlah orang yang terpapar *Covid-19* di Indonesia saat ini per tanggal 22 Agustus 2022 mencapai 6.318.857 orang, untuk kasus penularan *Covid-19* yang sembuh di Indonesia mencapai 6.112.658 orang, angka kematian mencapai 157.396 orang dan aktif sebanyak 48.803 orang. Sedangkan di Jawa Barat mencapai 1.157.533 orang, untuk kasus penularan *Covid-19* yang sembuh di Jawa Barat mencapai 1.124.258 orang, angka kematian mencapai 15.910 orang dan yang aktif sebanyak 17.365 orang sumber PIKOBAR (<https://pikobar.jabarprov.go.id/>). Berdasarkan data yang diperoleh angka positif *Covid-19* di Jawa Barat cukup banyak, sehingga perlu dilakukan pengelompokan angka penyebaran virus *Covid-*

19 di Jawa Barat berdasarkan jumlah orang yang terpapar, jumlah kematian, jumlah orang sembuh dan terkonfirmasi aktif yang nantinya bisa mengetahui Kabupaten atau kota yang terpapar *covid-19* dengan parameter *cluster* rendah, sedang dan tinggi. Proses pengelompokan orang yang terpapar *Covid-19* berdasarkan wilayah Kabupaten/Kota dengan cara membandingkan metode atau algoritma *Clustering* untuk mendapatkan hasil yang efisien dengan melakukan beberapa evaluasi pada algoritma tersebut.

Beberapa penelitian terkait dengan *Clustering* pada *Covid-19* adalah sebagai berikut. Penerapan Data Mining dalam Mengelompokkan Jumlah Kematian Penderita *Covid-19* Berdasarkan Negara di Benua Asia (Noviyanto, 2020). Hasil dari penelitian tersebut menggunakan metode *K-Means Clustering* menghasilkan 3 *cluster*, 4 negara dengan *cluster* tingkat tinggi yaitu: Turki, Iran, India dan China, *Cluster* tingkat sedang sebanyak 4 negara yaitu: Pakistan, Indonesia, Jepang dan Filipina serta *cluster* rendah adalah 41 negara lainnya. Adapun penelitian terkait yang membahas tentang penyebaran *Covid-19* di Jawa Barat oleh (Ramadanti and Muslih, 2021) dengan judul “Analisis Persebaran Kasus *Covid-19* Di Jawa Barat Menggunakan Metode *K-Means Clustering*” didapatkan hasil sebanyak 3 *cluster* yaitu tinggi, sedang dan rendah, yang termasuk ke dalam *cluster* tinggi terdiri dari 2 kabupaten/kota, Selanjutnya yang tergolong ke dalam *cluster* sedang terdapat 6 kabupaten/kota dan terdapat 19 kabupaten/kota lainnya. Selanjutnya ada penelitian tentang Komparasi Algoritma *K-Means* dan *K-Medoids* Untuk Pengelompokkan Penyebaran *Covid-19* di Indonesia, hasil dari *cluster* yang dibentuk dengan algoritma *K-Means* adalah *cluster* 1 memiliki 2 wilayah, sedangkan *cluster* 2

memiliki 32 wilayah. Hasil dari *cluster* yang dibentuk dengan menggunakan algoritma *K-Medoids* adalah *cluster* 1 memiliki 4 wilayah, sedangkan *cluster* 2 memiliki 30 wilayah. Hasil *cluster* dievaluasi menggunakan nilai koefisien shading dengan metode *K-Means* sebesar 0,207. Sedangkan nilai koefisien siluet dengan metode *K-Medoids* adalah 0,347 hal ini menunjukkan bahwa metode *K-Medoids* menghasilkan nilai SC yang lebih tinggi dibandingkan dengan *K-Means*, sehingga *K-Medoids* dapat memberikan hasil *Clustering* yang lebih baik. (Fira, Rozikin and Garno, 2021).

Berdasarkan penelitian (Fira, Rozikin and Garno, 2021) yang berjudul “Komparasi Algoritma *K-Means* dan *K-Medoids* Untuk Pengelompokan Penyebaran *Covid-19* di Indonesia” didapatkan hasil bahwa algoritma *K-Medoids* merupakan algoritma yang terbaik dibandingkan dengan algoritma *k-means* karena memiliki nilai *Silhouette Coefficient* paling optimal, adapun penelitian oleh (Farissa, Mayasari and Umidah, 2021) dengan judul “Perbandingan Algoritma *K-Means* dan *K-Medoids* Untuk Pengelompokan Data Obat dengan *Silhouette Coefficient*” didapatkan hasil bahwa algoritma *k-means* merupakan algoritma yang paling optimal dibandingkan dengan *k-medoids* karena memiliki nilai *Silhouette Coefficient* paling optimal, dan penelitian oleh (Rohmah, Rini and Utami, 2020) yang berjudul “Zonasi Daerah Terdampak Bencana Angin Puting Beliung Menggunakan *K-Means Clustering*” menggunakan evaluasi *Davies-Bouldin Index*, *Purity* dan *Silhouette Coefficient* dalam proses *clustering* menggunakan metode *k-means*. Berdasarkan pemaparan tersebut sehingga akan dilakukan penelitian pengelompokan jumlah positif *Covid-19* dengan membandingkan 2 metode

algoritma *Clustering* yaitu algoritma *K-Means* dan *K-Medoids* untuk mendapatkan nilai akurasi yang paling optimal dengan menggunakan 3 metode evaluasi yaitu *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient*, DBI nilai akurasi yang paling tepat mendekati atau sama dengan 0 sedangkan *Purity* dan *Silhouette Coefficient* nilai akurasi terbaiknya mendekati atau sama dengan 1. Berdasarkan uraian diatas penelitian ini berjudul “Perbandingan Algoritma *K-Means* Dan *K-Medoids* Untuk Pengelompokan Wilayah Penyebaran Jumlah Positif *Covid-19* Di Jawa Barat Dengan Analisis *Davies-Bouldin Index*, *Purity* dan *Silhouette Coefficient*”.

## 1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang, maka rumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana hasil *Clustering* menggunakan algoritma *K-Means* dan *K-Medoids* pada jumlah positif *Covid-19* di Jawa Barat?
2. Bagaimana mengukur akurasi hasil *cluster* jumlah positif *Covid-19* di Jawa Barat yang dievaluasi menggunakan *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient* pada algoritma *K-Means* dan *K-Medoids*?
3. Algoritma *Clustering* manakah yang memiliki nilai akurasi terbaik dari hasil evaluasi menggunakan *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient*?

### 1.3 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Mengetahui hasil *Clustering* menggunakan algoritma *K-Means* dan *K-Medoids* pada jumlah positif *Covid-19* di Jawa Barat.
2. Mengetahui hasil akurasi *cluster* jumlah positif *Covid-19* di Jawa Barat yang dievaluasi menggunakan *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient* pada Algoritma *K-Means* dan *K-Medoids*.
3. Mengetahui algoritma *Clustering* yang memiliki nilai akurasi terbaik menggunakan *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient*.

### 1.4 Manfaat Penelitian

Manfaat dari penelitian ini adalah:

1. Bagi Penulis

Dapat menambah pengetahuan dan wawasan serta lebih memahami mengenai metode *Clustering* menggunakan metode atau algoritma *K-Means* dan *K-Medoids*, perbandingan algoritma *K-Means* dan *K-Medoids* untuk mengukur nilai akurasi dengan evaluasi *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient*.

2. Bagi Pembaca

Dapat menambah pengetahuan baru mengenai metode atau algoritma *Clustering K-Means* dan *K-Medoids* dan pengukuran nilai akurasi dengan evaluasi *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient* yang dapat digunakan serta memberikan informasi bagi pihak yang membutuhkan.

### 1.5 Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Metode yang digunakan dalam *Clustering* atau pengelompokan jumlah positif *Covid-19* di Jawa Barat adalah algoritma *K-Means* dan *K-Medoids*.
2. Metode pengukuran nilai akurasi menggunakan evaluasi *Davies-Bouldin Index* (DBI), *Purity* dan *Silhouette Coefficient*.
3. Data yang digunakan yaitu data *Covid-19* di Jawa Barat dari 02 Maret 2020 sampai dengan 30 Juni 2022 yang diambil dari Open Data Jawa Barat Satuan Tugas Penanganan *Corona Virus Disease 2019* (SATGAS *COVID-19*).
4. Data yang diolah jumlah positif (Aktif, Sembuh dan Meninggal) *Covid-19* di Jawa Barat.