



Analisis Data Perkembangan Kasus Positif COVID-19 Di Jawa Barat Menggunakan Metode EDA (Exploratory Data Analysis)

Amalia Ayu Wardhani*¹, Kemal Ade Sekarwati²

Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Gunadarma, ^{1,2}

maliadhani31@gmail.com ¹, ade@staff.gunadarma.ac.id ²

Informasi Artikel

Diterima :13-10-2022
Direview :30-10-2022
Diterbitkan :25-11-2022

Kata Kunci

COVID-19, Deltacron,
Data Analysis, EDA,
Omicron, Python,
Jupyter

Abstrak

COVID-19 cases that have afflicted the people of Indonesia, especially West Java, have experienced an increase and decrease in confirmed patients with the COVID-19 virus. This caused the Data and Information Center of the Ministry of Health of the Republic of Indonesia to make observations to inform the number of confirmed, recovered and death cases to the Indonesian people using visualizations in the form of graphs. This study aims to determine the number of spikes in certain categories every week, identify the top five data every day and week, and find out the dynamics of COVID-19 cases by making a comparative graph between the accumulation of active, recovered, and dead cases using the Exploratory Data Analysis (EDA) method. As a result of this study, visualization in the form of a bar chart gives an overview of daily and weekly cases in West Java that have confirmed cases, deaths, and recoveries. Visualization in the form of a line chart gives a comparative picture of each daily and weekly case.

1. Pendahuluan

COVID-19 adalah penyakit yang mengakibatkan terinfeksi sistem pernapasan yang disebabkan oleh virus Severe Acute Respiratory Syndrome 2 (SARS-CoV-2) yang dimulai dengan gejala ringan seperti flu, hingga pneumonia. Kasus ini pertama kali ditemukan pada bulan akhir Desember 2019 di Wuhan, China. Awal bulan Desember 2019 sampai sekarang, COVID19 memiliki varian yang beragam diantaranya Gamma, Beta, Alpha, Delta, dan Omicron. Indonesia hanya mengenal varian Delta dan Omicron yang baru bermunculan akhir-akhir ini. Dalam paragraf pendahuluan, penulis menyatakan latar belakang penelitian, identifikasi masalah, keterbatasan masalah penelitian, perumusan masalah perumusan, dan tinjauan pustaka dan penelitian sebelumnya (referensi diambil dari artikel jurnal atau publikasi ilmiah maksimal Lima tahun kebelakang). Penulis juga menjelaskan kebaruan penelitian, tujuan penelitian dan signifikansinya.

Pada era yang sudah mutakhir ini, Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia melakukan pengamatan dengan menggunakan data tiap provinsi yang nantinya akan dicatat dalam big data. Pengembangan dari sistem basis data pada umumnya disebut big data. Teknologi basis data merupakan kumpulan data yang disimpan di dalam komputer secara sistematis sehingga dapat diperiksa menggunakan program komputer untuk memperoleh informasi dari basis data tersebut (Fikri, 2019). Exploratory Data Analysis atau yang dikenal dengan EDA merupakan salah satu strategi untuk melakukan analisis data, erat kaitannya dengan era big data saat ini.

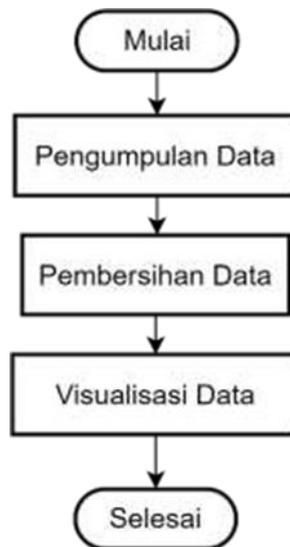
Sudah banyak peneliti yang meneliti tentang eksplorasi dan analisis data dengan menggunakan metode EDA seperti yang ditulis oleh Prismahardi Aji Riyantoko, Kartika Maulida Hindrayani, Tresna Maulana Fahrudin, dan I Gede Susrama Mas Diyasa pada tahun 2020 dalam jurnal yang berjudul “Exploratory Data Analysis pada Kasus COVID-19 di Indonesia Menggunakan HiveQL dan Hadoop Environment”. Pada penelitian tersebut menggunakan basis data terdistribusi yang bernama Hadoop Distributed File System (HDFS) dan jika data sudah tersimpan dapat dilakukan query menggunakan Hive Query Language (HiveQL).

Selain itu, penelitian yang ditulis oleh Abi Vegari dan Setia Budi pada tahun 2020 dengan judul “Implementasi Exploratory Data Analysis Pada Dataset Video Trending Harian YouTube” yang juga menggunakan metode EDA di dalam penelitiannya.

Berdasarkan latar belakang di atas, tertarik untuk melakukan penelitian mengenai “Analisis Data pada Perkembangan Kasus Positif COVID-19 di Jawa Barat Menggunakan Metode EDA (Exploratory Data Analysis)”. Data yang digunakan adalah data wilayah Jawa Barat karena Jawa Barat sempat menjadi penyumbang terbanyak kasus harian COVID-19. Menurut keterangan Kementerian Kesehatan yang dilansir oleh Media Indonesia pada tanggal 7 Maret 2022 mengatakan bahwa ada 4.354 kasus penularan lokal dan 14 kasus dari pelaku perjalanan luar negeri (PPLN).

2. Metode Penelitian

Tahapan penelitian yang digunakan dalam eksplorasi data COVID-19 pada Provinsi Jawa Barat menggunakan metode EDA yang terdiri dari pengumpulan data, pembersihan data dan visualisasi data. Berikut tahapan yang digunakan seperti pada Gambar 1.



Gambar 1. Metode Penelitian

Langkah-langkah dalam penelitian sebagai berikut:

1. Pengumpulan data. Pengumpulan data dilakukan dengan cara mengambil data melalui dokumen elektronik dari institusi Open Data Jabar dengan format .csv. yang berisi nama kabupaten/kota, pasien yang terkonfirmasi sembuh, aktif dan meninggal.
2. Pembersihan data. Pada tahap ini, kolom yang tidak digunakan dihilangkan. Pada data yang tertera, tidak ada kolom yang dihilangkan karena kolom tersebut seluruhnya

digunakan seperti kolom terkonfirmasi, meninggal, dan sembuh dari tanggal 1 Maret 2022 - 4 April 2022.

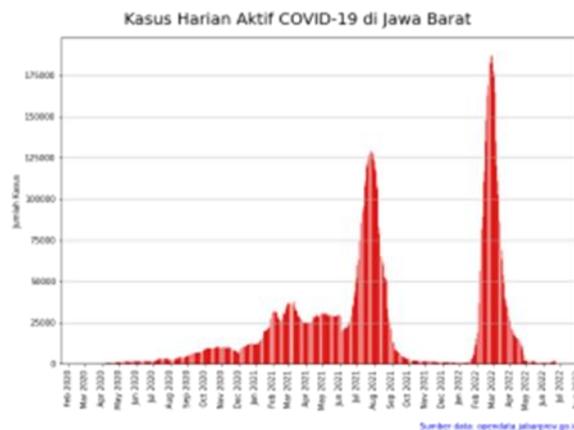
3. Visualisasi data. Setelah data dibersihkan, data akan divisualisasikan sesuai dengan yang dibutuhkan dalam penelitian. Tujuan dalam visualisasi data ini adalah untuk meringkas dan memperoleh pengetahuan dengan menggunakan grafik. Visualisasi berupa: 1. Grafik harian kasus aktif, meninggal, dan sembuh. 2. Grafik komparasi antara akumulasi kasus aktif, meninggal, dan sembuh. 3. Grafik fluktuasi pertambahan kasus harian. Tahap hasil analisis, menggunakan teori Kaplan dan Norton untuk menentukan hasil akhir penelitian rata-rata tingkat kepuasan pengguna.

3. Hasil dan Pembahasan

1. Membuat Grafik Harian Kasus Aktif, Meninggal, dan Sembuh

Dalam pembuatan grafik harian kasus aktif, meninggal, dan sembuh menggunakan grafik berupa bar chart dengan metode EDA. Berikut output:.

- Grafik Kasus Harian Aktif
- Dalam memvisualisasikan kasus harian aktif, grafik yang digunakan adalah jenis bar chart agar dapat dibaca lebih mudah oleh para masyarakat yang masih awam dalam memahami data.



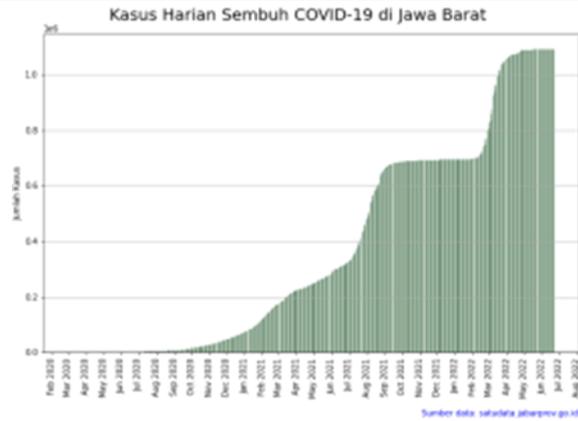
Grafik 1. Kasus Aktif COVID-19 di Jawa Barat

Hasil sumberdata opendata.jabarprov.go.id berdasarkan pada Grafik 1 memperlihatkan kasus pada tahun 2020 silam sempat mengalami kenaikan pada akhir Oktober. Pada saat COVID-19 baru saja masuk ke Indonesia tetapi angka kenaikan kasus aktif masih di bawah 12.500 kasus. Sempat terjadinya sedikit penurunan sekitar 10.000 kasus di akhir bulan November walaupun pada awal bulan Desember mengalami kenaikan kasus aktif COVID-19.

Pada tahun 2021 lagi-lagi mengalami kenaikan kasus yang jika dibandingkan pada Oktober 2020, pada Maret 2021 mengalami kenaikan ± 37.500 kasus. Setelah kejadian di bulan Maret 2021, di bulan Juni 2021 mengalami penurunan mencapai kurang dari 25.000 kasus tetapi pada bulan Agustus 2021 mengalami lonjakan kenaikan yang dapat mencapai 125.000 lebih kasus aktif dan di akhir bulan Desember 2021 mengalami penurunan yang sangat drastis yang dapat mencapai ± 100 kasus.

Di tahun 2022 di akhir bulan Juli mengalami lonjakan yang sangat pesat hingga mencapai 187.500 kasus dan semakin mereda hingga ± 100 kasus di akhir bulan Mei 2022.

Pada pembuatan grafik kasus harian sembuh dibutuhkan kolom tanggal dan juga kolom sembuh.

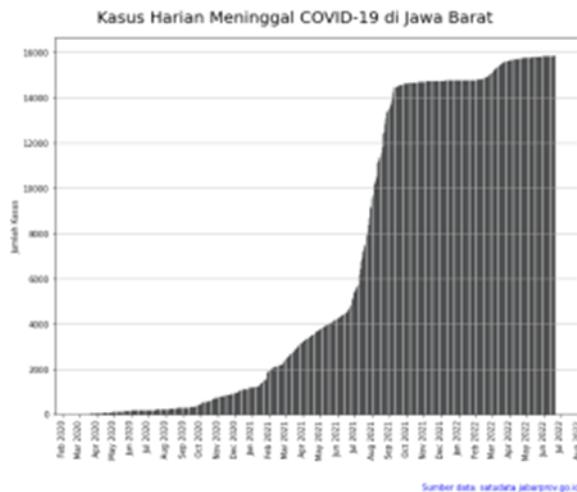


Grafik 2. Kasus Harian Sembuh COVID- 19 di Jawa Barat

Hasil dari sumber data yang dilansir opendata.jabarprov.go.id berdasarkan pada Grafik 2 memperlihatkan fluktuasi peningkatan kasus harian sembuh COVID- 19 mulai dari Juli 2020 hingga akhir bulan Juni 2022. Terlihat jelas bahwa kasus sembuhnya pasien yang terdampak COVID-19 di Provinsi Jawa Barat cenderung fluktuatif meningkat dari bulan ke bulan. Akhir Juni 2022 diketahui meningkat hingga 1.0×10^6 atau 1.000.000 kasus sembuh dari virus COVID-19.

- **Grafik Kasus Harian Meninggal**

Sama seperti grafik kasus harian sebelumnya, grafik kasus harian meninggal dibuat menggunakan bar chart.

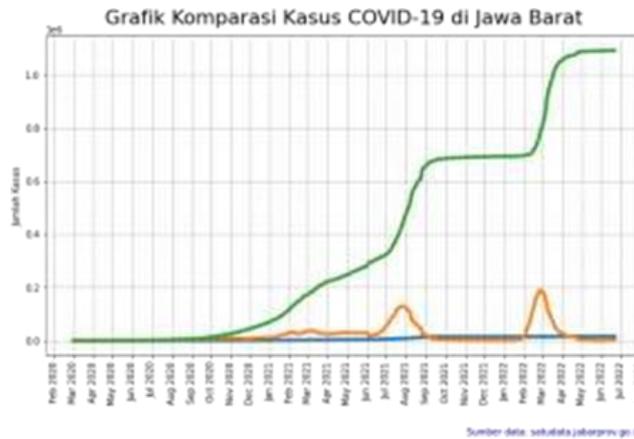


Grafik 3. Kasus Harian Meninggal COVID-19 di Jawa Barat

Hasil dari sumber data yang dilansir opendata.jabarprov.go.id berdasarkan pada Grafik 3 memperlihatkan fluktuasi peningkatan kasus harian meninggal COVID-19 mulai dari Maret 2020 hingga akhir bulan Juni 2022. Terlihat jelas bahwa kasus meninggalnya pasien yang terdampak COVID-19 di Provinsi Jawa Barat cenderung fluktuatif meningkat dari bulan ke bulan. Akhir Juni 2022 diketahui meningkat kurang dari 16.000 kasus sembuh dari virus COVID-19.

2. Membuat Grafik Komparasi Harian Antara Kasus Aktif, Meninggal, dan Sembuh

Pada pembuatan grafik komparasi harian antara kasus aktif, meninggal, dan sembuh menggunakan grafik berjenis line chart. Grafik yang digunakan terdiri dari gabungan grafik kasus aktif, kasus meninggal/kematian, dan kasus sembuh **Persamaan**



Grafik 1. Komparasi Kasus COVID-19 di Jawa Barat

Dengan menggunakan fitur tersebut, persamaan akan disajikan pada format seperti yang terlihat pada contoh di atas. Pastikan Anda tidak lupa menuliskan nomor yang diurutkan dengan format kesetaraan. Nomor persamaan ditulis dalam tanda kurung, yang diletakkan di sebelah kanannya.. Ketika Anda perlu menulis penjelasan yang mengacu pada persamaan, Anda dapat menulis dengan cara ini: dalam Persamaan (1) dan seterusnya. Persamaan ditulis dalam ukuran font normal, kecuali angka. Gunakan ukuran font 10 poin ke variabel, dan ukuran font 8 poin ke indeks.

Keterangan grafik sebagai berikut:

- Garis berwarna hijau untuk kasus sembuh.
- Garis berwarna jingga untuk kasus aktif.
- Garis berwarna biru untuk kasus meninggal.

Hasil komparasi pada 3 grafik yaitu kasus aktif, meninggal, dan sembuh menghasilkan meningkatnya kasus sembuh dari tahun ke tahun yang diperkirakan mencapai lebih dari 1.000.000 kasus.

3. Menampilkan Lima Data Teratas Untuk mengetahui data pertama tiap minggunya dengan mengatur kolom tanggal yang diaplikasikan dengan metode deret resampling mingguan dan agregasi jumlah. Ekstrak kolom tahun menggunakan metode .year lalu ubah nama kolom menjadi „tahun“ dan ekstrak data setiap minggu dalam satu tahun menggunakan metode .weekofyear lalu ubah nama kolom menjadi „minggu_ke“.

Lima data teratas covid_jabar_minggu: tahun minggu_ke jumlah

0	2020	9	2
1	2020	10	25
2	2020	11	63
3	2020	12	261
4	2020	13	640

Dapat disimpulkan pada output di atas bahwa di minggu ke-9 sebanyak 2 kasus terkonfirmasi virus, di minggu ke-10

sebanyak 25 kasus terkonfirmasi virus, pada minggu ke-11 sebanyak 63 kasus terkonfirmasi virus, pada minggu ke-12 sebanyak 261 kasus terkonfirmasi virus, dan pada minggu ke-13 sebanyak 640 kasus terkonfirmasi virus.

4. Menghitung Nilai Akumulasi Kasus

Aktif, Meninggal, dan Sembuh

Pada perhitungan kasus akumulasi aktif dibutuhkan kolom total_konfirmasi, sembuh dan meninggal untuk membuat tabel df_akumulasi dengan kolom akumulasi_aktif. Ketiga kolom tersebut lalu diakumulasikan menggunakan metode .cumsum. Pada tabel df_akumulasi untuk pembuatan kolom akumulasi_sembuh, kolom meninggal pada tabel df diakumulasikan menggunakan metode .cumsum. Tabel df_akumulasi pada pembuatan kolom akumulasi_meninggal menggunakan tabel df berkolum meninggal yang diakumulasikan menggunakan .cumsum

index	tanggal	akumulasi_aktif	akumulasi_sembuh	akumulasi_meninggal
836	2022-06-17 00:00:00	1643123	31175684	5649472
839	2022-06-18 00:00:00	16494556	312843075	5626321
840	2022-06-19 00:00:00	16496491	312838134	5641178
841	2022-06-20 00:00:00	16497712	315036882	5657034
842	2022-06-21 00:00:00	16499532	316121822	5672878

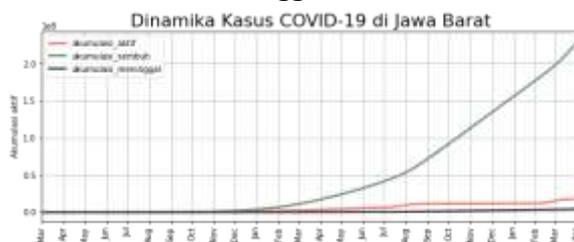
Gambar 2. Tabel Akumulasi

Terlihat pada 5 baris terakhir pada data yang menampilkan ekor dari seluruh data, di mana akumulasi kasus aktif meningkat tiap harinya. Dari tanggal 17 - 18 Juni 2022 meningkat sebanyak 1.433 kasus, 18 - 19 Juni 2022 meningkat sebanyak 1.535 kasus, 19 - 20 Juni 2022 meningkat sebanyak 1.621 kasus, dan 20 - 21 Juni 2022 meningkat sebanyak 1.820 kasus. Jika diperhatikan, adanya kenaikan kasus yang terjadi dari 20 - 21 Juni 2022 hampir mencapai 200 kasus dibanding hari-hari sebelumnya dan untuk akumulasi kasus sembuh dan kematian makin meningkat hari demi hari.

5. Menghitung Nilai Akumulasi Kasus

Aktif, Meninggal, dan Sembuh

Pada visualisasi data untuk komparasi antara akumulasi kasus aktif, meninggal, dan sembuh menggunakan tabel df_akumulasi supaya dapat membandingkan akumulasi kasus aktif, meninggal, dan sembuh.



Grafik 2. Komparasi antara Akumulasi Aktif, Meninggal, dan Sembuh

Tampilan pada Grafik 5 menunjukkan bahwa akumulasi kasus sembuh meningkat dari tahun ke tahun. Akumulasi kasus aktif berdasarkan pada Grafik 5 menunjukkan bahwa masyarakat harus tetap waspada menggunakan masker dan menjaga kebersihan badan dan lingkungan guna menjaga kelonjangan yang disebabkan naiknya akumulasi kasus

aktif. Dan beruntungnya, pada akumulasi meninggal yang terdapat di Grafik 5 hanya sedikit dari tahun ke tahun.

4. KESIMPULAN SARAN

Analisis data perkembangan harian kasus positif COVID-19 di Jawa Barat menggunakan metode EDA (Exploratory Data Analysis) telah berhasil dibuat. Visualisasi dalam bentuk bar chart memberi gambaran mengenai kasus harian dan mingguan di Jawa Barat yang memiliki kasus terkonfirmasi, kematian, dan kesembuhan. Visualisasi dalam bentuk line chart memberi gambaran komparasi tiap kasus harian dan mingguan.

Penelitian ini hanya menggunakan *bar chart* dan *line chart*. Penelitian ini diharapkan dapat mengembangkan model visualisasi data seperti *pie chart*, *line plot*, *scatter plot*, dan *bar plot*. Kekurangan metode EDA dalam kasus yang diangkat tidak dapat memprediksi adanya kenaikan atau penurunan angka kasus COVID-19. Saran dari penelitian selanjutnya metode EDA dapat dikembangkan secara bersamaan dengan metode regresi linier dan deskripsi agar dapat memprediksi adanya kenaikan dan penurunan kasus COVID-19.

DAFTAR PUSTAKA

Hengki Wijaya. 2018. Analisis Data Kualitatif Ilmu Pendidikan Teologi. Makassar: Sekolah Tinggi Theologia Jaffray.

Indwiarti, Aniq Atiqi Rohmawati, Putu Harry Gunawan. 2021. Pengantar Statistika Eksplorasi Data dengan Ms. Excel dan R. Yogyakarta: KBM Indonesia.

Robert Kurniawan. 2019. Cara Mudah Belajar Statistik Analisis Data & Eksplorasi. Jakarta: Prenada Media.

URL : <https://www.alodokter.com/covid-19> . Diakses pada 6 April 2022

URL: <http://www.bccdc.ca/health-info/diseases-conditions/covid-19/about-covid-19/variants>. Diakses pada 6 April 2022

URL : <https://glints.com/id/lowongan/jupyter-adalah/#.YvTMUdBy00> . Diakses pada 6 April 2022
URL : <https://medium.com/purwadhikaconnect/analisa-data-kejahatan-di-boston-menggunakan-python-55c42934b739> . Diakses pada 6 April 2022

URL : <https://medium.com/@chauhanhrithik/exploratory-data-analysis-7f359874d615> . Diakses pada 6 April 2022

URL : <https://www.pythonindo.com/sejarah-python/> . Diakses pada 6 April 2022