

Visualisasi Data Dari Dataset COVID-19 Menggunakan Pemrograman Python

(Center, Times New Roman 18 **not Bold**, maksimum. 12 kata Bahasa Indonesia atau 10 words in English)

Judul Naskah Bahasa Inggris (Font. 12 Not Bold, Italic)

Yanuangga Galahartlambang¹, Titik Khotiah², Jumain³

^{1,2,3} Teknologi Informasi, Fakultas Teknik,

Institut Teknologi dan Bisnis Ahmad Dahlan Lamongan, Indonesia

e-mail: ¹yanuangga@gmail.com, ²titikaye@gmail.com, ³jumain.dj@gmail.com

Abstrak

Dunia telah menyaksikan pada tahun 2020 wabah global SARS-CoV-2 yang belum pernah terjadi sebelumnya, jenis virus corona baru, yang menyebabkan pandemi COVID-19, dan secara radikal mengubah kehidupan dan pola kerja kita. Penelitian ini bertujuan untuk menghasilkan visualisasi atau gambaran data berdasarkan penyebaran kasus Covid-19 di Indonesia. Dataset yang digunakan pada penelitian ini menggunakan data kasus covid-19 di provinsi DKI Jakarta dalam bulan februari 2021 untuk menyajikan visualisasi data yang beragam untuk membantu pihak berwenang dalam mengambil keputusan yang tepat untuk menangani masalah yang belum pernah terjadi sebelumnya. Python dan perpustakaan bersama dengan platform Google Colab digunakan untuk mendapatkan hasilnya. Teknik terbaik dan kombinasi modul/pustaka digunakan untuk menyajikan informasi terkait COVID-19.

Kata kunci— Covid-19, visualisasi data, data science, pandemi

Abstract

The world has witnessed in 2020 an unprecedented global outbreak of SARS-CoV-2, a new type of coronavirus, which caused the COVID-19 pandemic, and radically changed our lives and work patterns. This study aims to produce a visualization or picture of data based on the spread of Covid-19 cases in Indonesia. The dataset used in this study uses data on COVID-19 cases in the DKI Jakarta province in February 2021 to present a variety of data visualizations to assist the authorities in making the right decisions to deal with this unprecedented problem. Python and its libraries along with the Google Colab platform were used to obtain the results. The best techniques and a combination of modules/libraries are used to present information related to COVID-19.

Keywords— Covid-19, data visualization, data science, pandemic

PENDAHULUAN

Penyakit corona virus 2019 (COVID-19) telah memicu pandemi di seluruh dunia (Kamal et al., 2020). Penyakit ini disebabkan oleh *Severe Acute Respiratory Syndrome coronavirus-2* (SARS-CoV-2). Seperti kita semua ketahui, pandemi COVID-19 telah berdampak pada banyak negara termasuk Indonesia adalah salah satunya negara dengan jumlah kasus yang tinggi. Penyebaran virus corona terjadi secara radikal, banyak sektor kegiatan bisnis terhenti hingga kebijakan mengenai pembatasan aktivitas sosial dilakukan oleh pemerintah, yang berdampak pada depresi ekonomi sehingga menyebabkan terpuruknya negara diberbagai sektor usaha. Penelitian ini menyajikan analisis dan visualisasi data penyebaran COVID-19 dari 1 maret 2020 hingga 31 maret 2020 mengambil dataset dari provinsi DKI Jakarta.

Informasi Artikel:

Submitted: bulan 20xx, **Accepted:** bulan 20xx, **Published:** Mei / November 20xx

ISSN: 2685-4902 (media online), Website: <http://jurnal.umus.ac.id/index.php/intech>

METODE PENELITIAN

Visualisasi data adalah representasi grafis dari data. Hal ini penting dalam menganalisis data karena akan membantu untuk pengguna untuk memahami data seperti yang ditampilkan dalam infografis. Orang cenderung melihat elemen visual sebagai informasi sederhana dan lebih menarik daripada teks data yang rumit, sehingga tidak memerlukan waktu lebih lama untuk menerjemahkan informasi dalam visualisasi data. Visualisasi data pada penelitian ini akan menyajikan jumlah kasus yang terjadi setiap kelurahan di provinsi DKI Jakarta.

Tujuan utama visualisasi data virus Corona adalah untuk mengkomunikasikan informasi dengan jelas dan efektif menggunakan presentasi grafis yang berbeda. visualisasi berguna media untuk memeriksa, memahami, dan mentransmisikan informasi karena memiliki beberapa kemungkinan kegunaan dalam domain.

Bahasa pemrograman Python sebagai salah satu pemrograman terbaik untuk menangani visualisasi data karena memiliki dukungan komunitas yang aktif dan memiliki banyak pustaka bantu terutama dalam hal komputasi cerdas. Bahasa pemrograman mampu menangani data dalam jumlah besar dengan performa yang baik.

2.1 Pustaka Visualisasi Data dengan python

2.1.1 Matplotlib

Matplotlib adalah salah satu paket pustaka Python yang paling populer dalam hal visualisasi data. Paket pustaka ini dapat menyajikan model visual plot 2D dari data dalam bentuk array. Pustaka ini juga menyediakan API (*Application Programming Interface*) yang dapat digunakan dalam aplikasi bantu Python GUI seperti *PyQt*, *WxPython* dan *Tkinter*. Selain itu juga dapat digunakan dalam shell Python dan *IPython*, *notebook Jupyter* dan *web server*.

2.1.2 Seaborn

Pustaka ini dibuat dari paket pustaka *Matplotlib*. Yang memungkinkan pengembang aplikasi untuk membuat visualisasi data variatif sesuai dengan kebutuhan (seperti memetakan warna ke dalam variabel). *Seaborn* lebih terintegrasi untuk bekerja dengan *Pandas DataFrames*.

2.1.3 Plotly

Plotly.js adalah data *JavaScript* deklaratif pustaka visualisasi yang dibangun di atas *D3* dan *WebGL* yang mendukung berbagai statistik, ilmiah, keuangan, geografis, dan visualisasi 3 dimensi.

2.1.4. Folium

Pustaka Ini mampu menampilkan visualisasi data dalam bentuk peta dari sekumpulan data.

2.2 Corona Virus Disease (Covid-19)

Virus corona atau *severe acute respiratory syndrome coronavirus 2 (SARSCoV-2)* adalah virus yang menyerang sistem pernapasan manusia. Penyakit ini dikarenakan infeksi virus yang disebut Covid-19. Virus corona bisa menyebabkan gangguan ringan pada sistem pernapasan, infeksi paru-paru yang berat, hingga kematian. *Severe acute respiratory syndrome corona virus 2 (SARS-CoV-2)* yang lebih dikenal dengan nama virus corona adalah jenis baru dari corona virus yang menular ke manusia. Virus ini bisa menyerang siapa saja, baik bayi, anak-anak, orang dewasa, lansia, ibu hamil, maupun ibu menyusui (Handayani, 2020). Corona virus adalah kumpulan virus yang bisa menginfeksi sistem pernapasan (Kemenkes, 2020).

2.2.1 Manifestasi klinis

Gejala klinis umum yang terjadi pada pasien Covid19, diantaranya yaitu demam, batuk kering, *dispnea*, *fatigue*, nyeri otot, dan sakit kepala (Lapostolle dkk, 2020). Berdasarkan penelitian yang dilakukan oleh Huang dkk (2020), gejala klinis yang paling sering terjadi pada pasien Covid19 yaitu demam (98%), batuk (76%), dan myalgia atau kelemahan (44%). Gejala lain yang terdapat pada pasien, namun tidak begitu sering ditemukan yaitu produksi sputum (28%), sakit kepala 8%, batuk darah 5%, dan diare 3%, sebanyak 55% dari pasien yang diteliti mengalami *dispnea*.

Gejala klinis yang melibatkan saluran pencernaan juga dilaporkan oleh Kumar dkk (2020). Sakit abdominal merupakan indikator keparahan pasien dengan infeksi Covid19. Sebanyak 2,7% pasien mengalami sakit abdominal, 7,8% pasien mengalami diare, 5,6% pasien mengalami mual dan/atau muntah.

Computerised Tomographytoraks (CT toraks) pada pasien dengan Covid19 pada umumnya memperlihatkan opasifikasi ground-glass dengan atau tanpa gabungan abnormalitas. CT toraks mengalami abnormalitas bilateral, distribusi perifer, dan melibatkan lobus bawah. Penebalan *pleural*, *efusi pleura*, dan *limfadenopati* merupakan penemuan yang jarang didapatkan (Gennaro dkk, 2020). Individu yang terinfeksi namun tanpa gejala dapat menjadi sumber penularan SARS-CoV-2 dan beberapa diantaranya mengalami progres yang cepat, bahkan dapat berakhir pada ARDS dengan case fatality rate tinggi. Penelitian yang dilakukan oleh Meng dkk tahun 2020 menunjukkan bahwa dari 58 pasien tanpa gejala yang dites positif Covid19 pada saat masuk RS, seluruhnya memiliki gambaran CT-Scan toraks abnormal. Penemuan tersebut berupa gambaran opasitas ground-glass dengan distribusi perifer, lokasi unilateral, dan paling sering mengenai dua lobus paru. Setelah follow up dalam jangka waktu singkat, 27,6% pasien yang sebelumnya asimtomatik mulai menunjukkan gejala berupa demam, batuk, dan *fatigue*.

2.2.2 Penularan

Penularan ini terjadi umumnya melalui droplet dan kontak dengan virus kemudian virus dapat masuk ke dalam mukosa yang terbuka. Suatu analisis mencoba mengukur laju penularan berdasarkan masa inkubasi, gejala dan durasi antara gejala dengan pasien yang diisolasi. Analisis tersebut mendapatkan hasil penularan dari 1 pasien ke sekitar 3 orang di sekitarnya, tetapi kemungkinan penularan di masa inkubasi menyebabkan masa kontak pasien ke orang sekitar lebih lama sehingga risiko jumlah kontak tertular dari 1 pasien mungkin dapat lebih besar (Handayani, 2020).

2.2.3 Diagnosis

2.2.3.1 Kasus Suspek

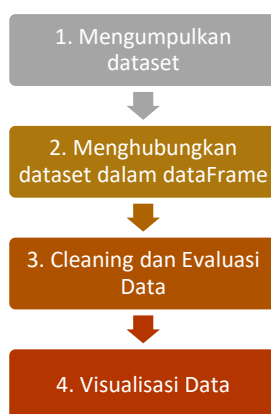
Seseorang yang memenuhi salah satu kriteria klinis dan salah satu kriteria epidemiologis dan kriterian klinis. Kriteria Klinis: demam akut : $\geq 38^{\circ}\text{C}$ /riwayat demam dan batuk, atau terdapat 3 atau lebih gejala/tanda akut berikut: demam/riwayat demam, batuk, kelelahan (*fatigue*), sakit kepala, *myalgia*, nyeri tenggorokan, *coryza*/ pilek/ hidung tersumbat, sesak nafas, *anoreksia*/mual/muntah, diare, penurunan kesadaran. Seseorang dengan infeksi saluran pernafasan atas berat. Seseorang dengan gejala akut anosmia (hilangnya kemampuan indra penciuman) atau ageusia (hilangnya kemampuan indra perasa) dengan tidak ada penyebab lain yang dapat diidentifikasi (Handayani, 2020)

2.2.3.2 Kasus Probable

Kasus suspek yang meninggal dengan gambaran klinis yang meyakinkan Covid 19 dan memiliki salah satu kriteria sebagai berikut: tidak dilakukan pemeriksaan laboratorium *RT-PCR*, hasil pemeriksaan laboratorium *RT-PCR* satu kali negatif dan tidak dilakukan pemeriksaan laboratorium *RT-PCR* yang kedua.

2.2.3.3 Kasus Konfirmasi

Seseorang yang dinyatakan positif terinfeksi virus Covid19 yang dibuktikan dengan pemeriksaan laboratorium *RT-PCR*. Kasus konfirmasi dibagi menjadi 2 yaitu Kasus konfirmasi dengan gejala (*simptomatik*) dan Kasus konfirmasi tanpa gejala (*asimptomatik*).



Gambar 1 Kerangka kerja Penelitian

2.3 Kerangka Kerja Penelitian

Secara garis besar kerangka kerja pada penelitian ini adalah :

2.3.1 Mengumpulkan dataset : Mengumpulkan data terkait informasi kasus covid-19 khususnya untuk daerah provinsi DKI Jakarta dengan rentang waktu tertentu

2.3.2 Menghubungkan dataset kedalam dataframe : Dataset yang telah terkumpul, akan di muat ke dalam *dataframe* menggunakan pustaka pemrograman python sebagai informasi untuk menampilkan visualisasi data

2.3.3 *Cleaning* dan Evaluasi data : Pada tahap ini dataset yang telah dimuat di *dataframe* dievaluasi dan dibersihkan dari data kosong (*missed value*) agar tidak memicu terjadinya kegagalan dalam menampilkan informasi

2.3.4 Visualisasi data : Dataframe yang sudah dievaluasi akan ditampilkan menggunakan pustaka pemrograman python dengan berbagai model visualiasi data untuk memudahkan penyampaian informasi ke pengguna

HASIL DAN PEMBAHASAN

Dataset yang telah diperoleh disimpan dalam format csv (comma delimited) yang terdiri dari 44 baris dan 14 kolom, dengan penjelasan atribut seperti pada tabel 1.

Tabel 1. Atribut tabel dataset kasus covid-19 provinsi DKI Jakarta

#	Kolom	Keterangan
1	Id_kel	Kode kelurahan
2	nama_kota	Nama kota / kabupaten
3	nama_kecamatan	Nama kecamatan
4	odp	Orang dalam pemantauan
5	proses_pemantauan	Orang dalam proses pemantauan
6	selesai_pemantauan	Orang selesi proses pemantauan
7	Pdp	Pasien dalam pemantauan
8	masih_dirawat	Pasien dengan status masih dirawat
9	pulang_dan_sehat	Pasien sudah pulang dan dinyatakan sehat
10	Positif	Pasien dinyatakan positif covid-19
11	Dirawat	Pasien sedang dalam perawatan
12	Sembuh	Pasien dinyatakan telah sembuh dari covid-19
13	Meninggal	Pasien meninggal karena covid-19
14	self_isolation	Pasien sedang isolasi mandiri

Setelah dataset dikumpulkan langkah berikutnya menghubungkan dataset kedalam dataframe python, tujuannya untuk proses analisa dan visualisasi data. Gambar 2 merupakan hasil dataframe dari kasus covid-19 provinsi DKI Jakarta.

id_kel	nama_kota	nama_kecamatan	odp	proses_pemantauan	selesai_pemantauan	pdp	masih_dirawat	pulang_dan_sehat	positif	dirawat	sembuh	meninggal	self_isolation
3175061001	JAKARTA TIMUR	CAKUNG	19	5	14	22	15	7	12	9	0	2	1
3171051001	JAKARTA PUSAT	CEMPAKA PUTIH	17	4	13	11	9	2	5	2	1	0	2
3173011001	JAKARTA BARAT	CENGKARENG	28	8	20	10	7	3	9	6	0	2	1
3174061001	JAKARTA SELATAN	CILANDAK	37	8	29	41	24	17	22	10	3	3	6
3172041001	JAKARTA UTARA	CILINCING	34	10	24	14	11	3	3	1	0	2	0
3175101001	JAKARTA TIMUR	CIPAYUNG	25	8	17	13	12	1	6	5	1	0	0

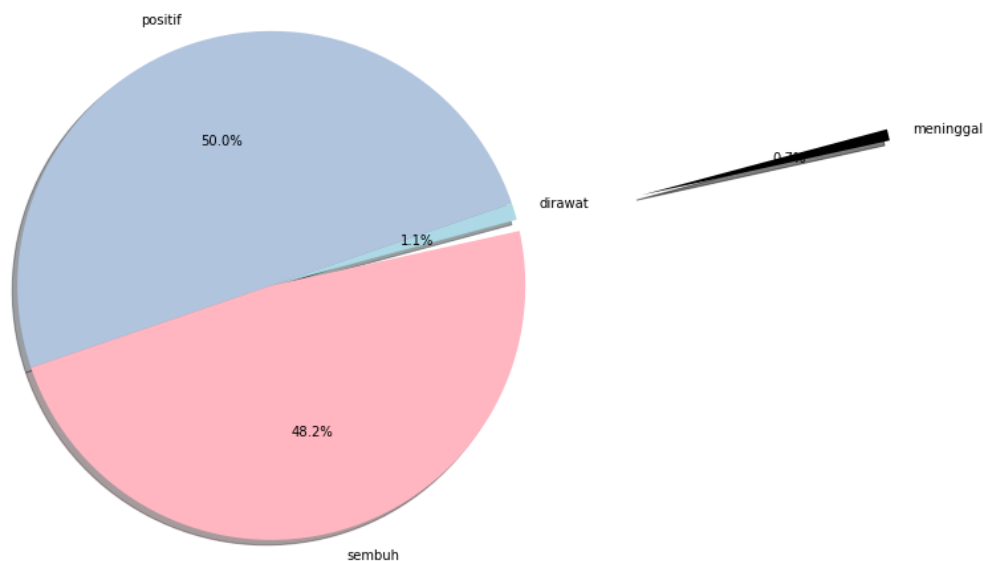
Gambar 2 Potongan Dataframe kasus covid-19 provinsi DKI Jakarta (maret 2020)

Dataset yang telah dihubungkan ke dalam dataframe kemudian dilakukan evaluasi dan analisa. Pada dataframe gambar 2 terdapat kolom yang dapat diabaikan atau tidak perlu ditampilkan sebagai informasi visual yaitu kolom id_kel (kode kelurahan). Kolom id_kel selanjutnya dihapus dari dataframe, sehingga menyisakan 13 kolom.

3.1 Visualisasi Data

3.1.1 Kasus positif, dirawat, sembuh dan meninggal DKI Jakarta

Hasil visualisasi kasus positif, dirawat, sembuh dan meninggal pada provinsi dapat dilihat

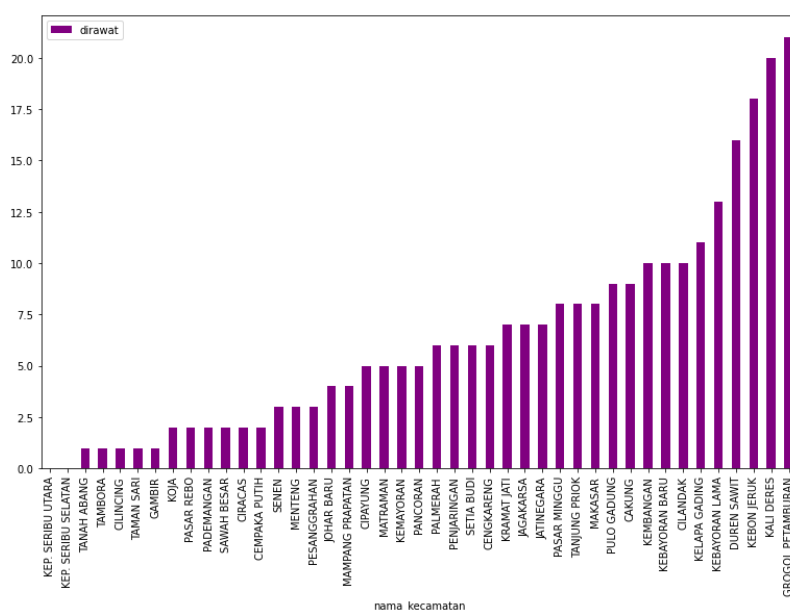


Gambar 3 Pie Chart Kasus positif, dirawat, sembuh dan meninggal

pada gambar 3. Dimana jumlah kasus pasien positif covid-19 sebanyak 50%, pasien berhasil sembuh sebanyak 48,2 %, pasien sedang dirawat sebanyak 1,1 % dan pasien covid-19 meninggal sebanyak 0,7 %.

3.1.2 Kasus pasien dirawat per-kecamatan

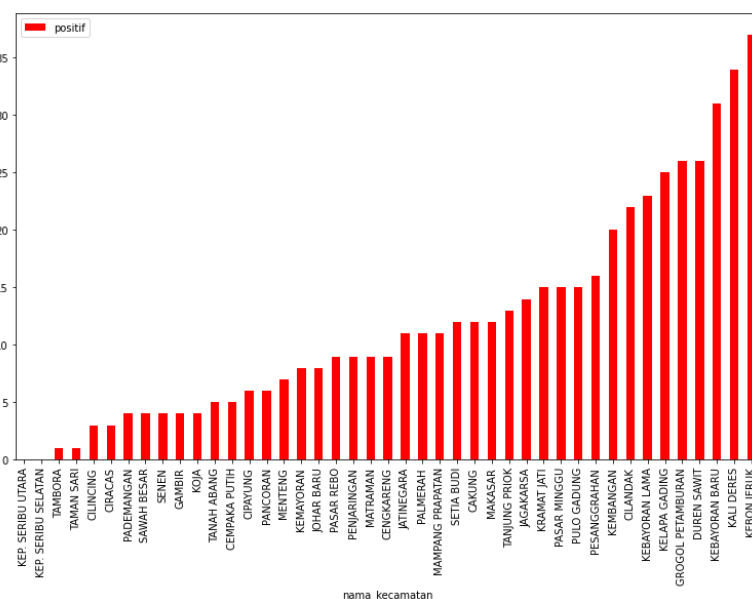
Hasil visualisasi kasus pasien sedang dirawat dengan tampilan data per-kecamatan di provinsi DKI Jakarta dapat dilihat pada gambar 4. Dari hasil visualisasi dapat disimpulkan kasus pasien dirawat paling banyak terjadi pada kecamatan Grogol Petamburan, sedangkan kasus paling sedikit terjadi di kecamatan kepulauan seribu utara dan selatan.



Gambar 4 Kasus Pasien dirawat per-kecamatan di DKI Jakarta

3.1.3 Kasus pasien positif per-kecamatan

Hasil visualisasi kasus pasien positif covid-19 dengan tampilan data per-kecamatan di provinsi DKI Jakarta dapat dilihat pada gambar 5. Dari hasil visualisasi dapat disimpulkan kasus

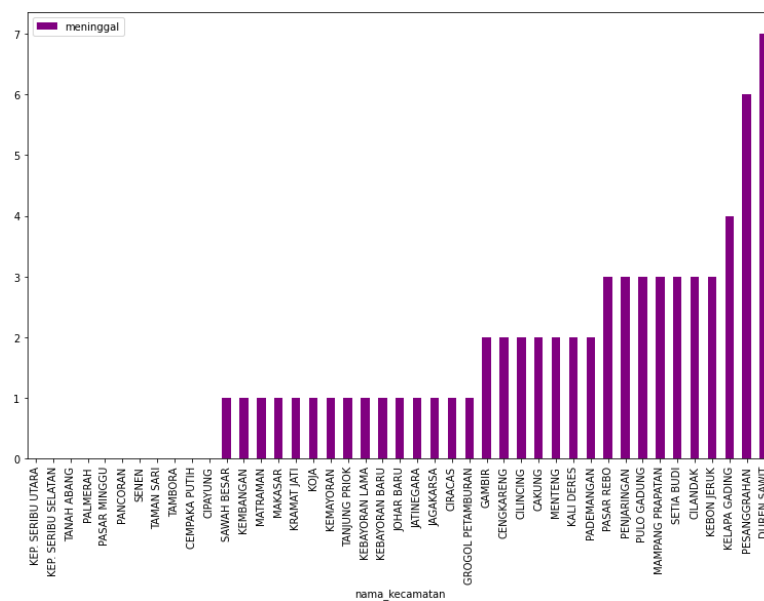


Gambar 5 Kasus Pasien positif covid-1 per-kecamatan di DKI Jakarta

pasien positif paling banyak terjadi pada kecamatan Kebon Jeruk, sedangkan kasus paling sedikit terjadi di kecamatan kepulauan seribu utara dan selatan.

3.1.4 Kasus pasien meninggal per-kecamatan

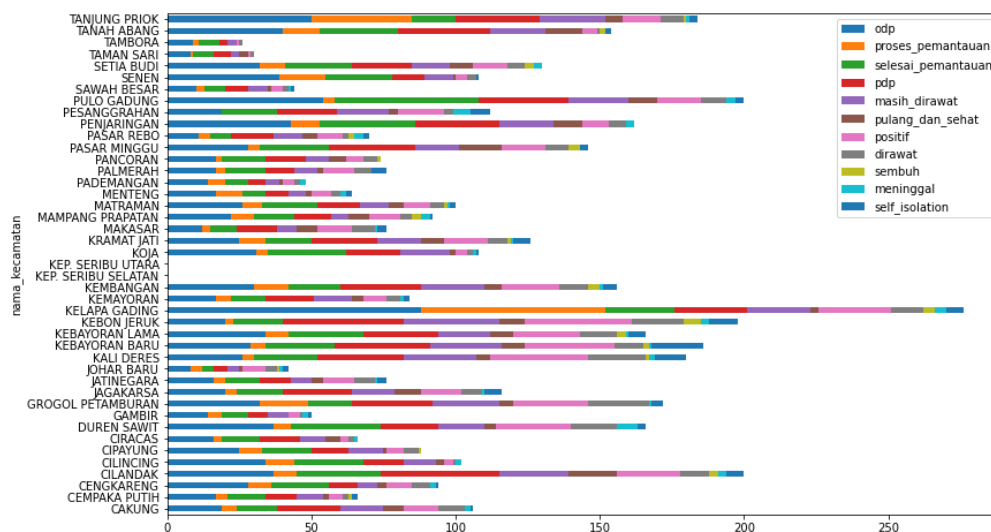
Hasil visualisasi kasus pasien positif covid-19 dengan tampilan data per-kecamatan di provinsi DKI Jakarta dapat dilihat pada gambar 6. Dari hasil visualisasi dapat disimpulkan kasus pasien meninggal paling banyak terjadi pada kecamatan Duren Sawit, sedangkan kasus paling sedikit terjadi di kecamatan kepulauan seribu utara, kep seribu selatan, tanah abang, pal merah, pasar minggu, pancoran, senen, taman sari, tambora, cempaka putih, cipayung.



Gambar 6 Kasus Pasien meninggal covid-1 per-kecamatan di DKI Jakarta

3.1.5 Seluruh Kasus yang terjadi per-kecamatan

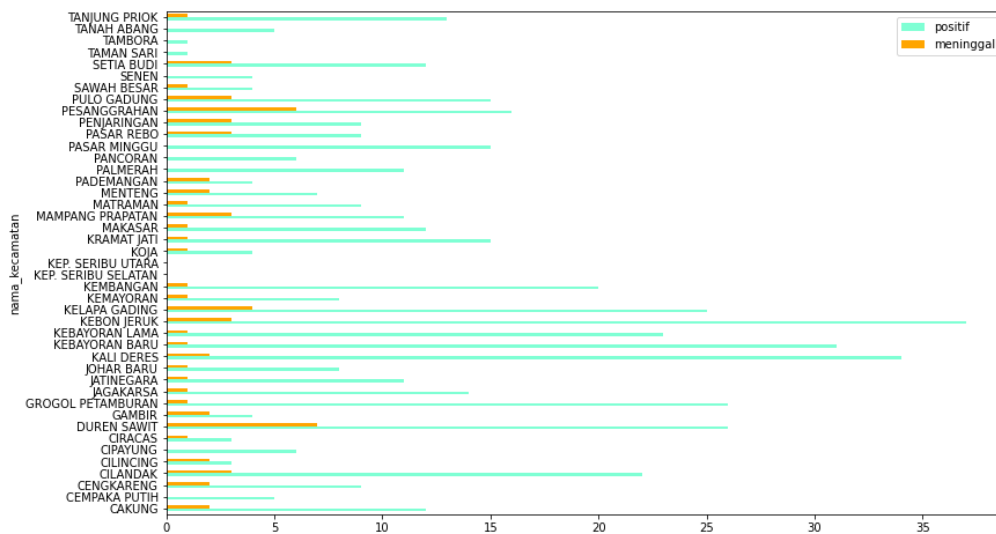
Hasil visualisasi seluruh kasus covid-19 yang terjadi, dengan tampilan data per-kecamatan di provinsi DKI Jakarta dapat dilihat pada gambar 7.



Gambar 7 Seluruh Kasus yang terjadi, per-kecamatan di DKI Jakarta

3.1.5 Kasus Positif dan Meninggal yang terjadi per-kecamatan

Hasil visualisasi kasus pasien positif dan meninggal karena covid-19 yang terjadi, dengan tampilan data per-kecamatan di provinsi DKI Jakarta dapat dilihat pada gambar 8.



Gambar 8 Kasus Positif dan Meninggal per-kecamatan di DKI Jakarta

KESIMPULAN

Kesimpulan yang dapat diambil dari penelitian ini adalah Visualisasi dalam bentuk laporan grafik dapat dengan mudah memahami persoalan, para pengambil keputusan dapat dengan cepat mengambil kebijakan dan pemrograman Python beserta pustaka pendukungnya mengoptimalkan permasalahan data yang besar dalam berbagai lingkungan baik pemerintah maupun manajemen bisnis.

DAFTAR PUSTAKA

- [1] Cesar Aguado-Cortés and Victor M Castaño. 2020. *Translational Knowledge Map of COVID-19*. arXiv preprint arXiv:2003.10434 (2020)
- [2] Sabber Ahamed and Manar Samad. 2020. *Information Mining for COVID-19 Research From a Large Volume of Scientific Literature*. arXiv:2004.02085
- [3] Mohamad Chahrour, et al. 2020. *A bibliometric analysis of Covid-19 research activity: A call for increased output*. Cureus 12, 3 (2020).
- [4] Iztok Fister Jr, Karin Fister, and Iztok Fister. 2020. *Discovering associations in COVID-19 related research papers*. arXiv preprint arXiv:2004.03397 (2020)
- [5] OpenData, Jakarta, Data Rekap Harian Kasus Covid19 Per Kelurahan di Provinsi DKI Jakarta Bulan Maret 2020, <https://data.jakarta.go.id/dataset/data-rekap-harian-kasus-covid19-per-kelurahan-di-provinsi-dki-jakarta-bulan-maret-2020>, diakses tanggal 22 Januari 2021 pukul 10.00
- [6] Milad Haghani, Michiel CJ Bliemer, Floris Goerlandt, and Jie Li. 2020. *The scientific literature on Coronaviruses, COVID-19 and its associated safety-related research dimensions: A scientometric analysis and scoping review*. Safety Science (2020), 104806
- [7] Pierre Le Bras, et al. 2018. *Improving User Confidence in Concept Maps: Exploring Data Driven Explanations*. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 3173978, 1–13. <https://doi.org/10.1145/3173574.3173978>
- [8] Coronaviridae Study Group of the International et al. 2020. *The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2*. Nature Microbiology (2020)
- [9] Stefano Padilla, et al. 2014. *Hot topics in CHI: trend maps for visualising research*. 815–824

- [10] Oberoi, A., & Chauhan, R. (2019). *Visualizing data using Matplotlib and Seaborn libraries in Python for data science*. International Journal of Scientific and Research Publications (IJSRP), 9(3), p8733. <https://doi.org/10.29322/ij srp.9.03.2019.p8733>
- [11] Comba, J. L. D. (2020). *Data Visualization for the Understanding of COVID-19*. Computing in Science Engineering, 22(6), 81–86. <https://doi.org/10.1109/MCSE.2020.3019834>
- [12] Jin YH, et al. 2020. *A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version)*. Military Medical Research.
- [13] Handayani, dkk. 2020. *Penyakit Virus Corona 2019*. Jurnal Respirologi Indonesia, 40(2), 119–129.
- [14] Susilo, A., dkk. 2020. *Coronavirus Disease 2019 : Tinjauan Literatur Terkini*. Jurnal Penyakit Dalam Indonesia, 7(1), 45–67
- [15] Lapostolle, et al. 2020. *Clinical Features of 1487 COVID - 19 Patients with Outpatient Management in the Greater Paris : the COVID - Call Study*. Internal and Emergency Medicine, (0123456789). <https://doi.org/10.1007/s11739-020-02379-z>
- [16] Huang, C., et al. 2020. *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China*. The Lancet. 24 jan 2020
- [17] Gennaro, F. Di, et al. 2020. *Coronavirus Diseases (COVID-19) Current Status and Future Perspectives : A Narrative Review*. International Journal of Environmental Research and Public HealthEnvironmental Research and Public Health, 17(2690), 1–11. <https://doi.org/10.3390/ijerph17082690>
- [18] Alfaridzi, M. A. ., & Premana, A. . (2021). *Aplikasi Berbasis WEB-Realtime Pemantauan Coronavirus Disease 2019 (COVID-19)*. Jurnal Ilmiah Intech : Information Technology Journal of UMUS, 3(01), 50–57. Retrieved from <http://jurnal.umus.ac.id/index.php/intech/article/view/416>