**LAPORAN TUGAS**

**ANALISIS DATA COVID 19 DATASET**

**MENGGUNAKAN POWER BI**



**DISUSUN OLEH :**

Moh. Abdan Syakura Putra (210411100011)

Dosen Pengampu : EKA MALA SARI ROCHMAN, S.Kom., M.Kom.

**PRODI TEKNIK INFORMATIKA**

**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS TRUNOJOYO MADURA**

**2023**

**COVID-19**

Coronavirus merupakan kelompok besar virus yang dapat menyebabkan penyakit pada manusia maupun hewan. Pada manusia, virus ini umumnya mengakibatkan infeksi saluran pernapasan, mulai dari gejala ringan seperti flu biasa hingga penyakit serius seperti Middle East Respiratory Syndrome (MERS) dan Severe Acute Respiratory Syndrome (SARS). Virus corona baru yang muncul di Wuhan, Cina pada Desember 2019 diberi nama Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), dan menyebabkan penyakit yang kita kenal sebagai Coronavirus Disease 2019 (COVID-19).

Gejala umum COVID-19 meliputi demam dengan suhu lebih dari 38°C, batuk kering, dan kesulitan bernapas. Jika seseorang memiliki riwayat perjalanan ke daerah terjangkit atau kontak dengan penderita COVID-19 dalam 14 hari sebelum munculnya gejala, maka akan dilakukan pemeriksaan lebih lanjut untuk konfirmasi diagnosis.

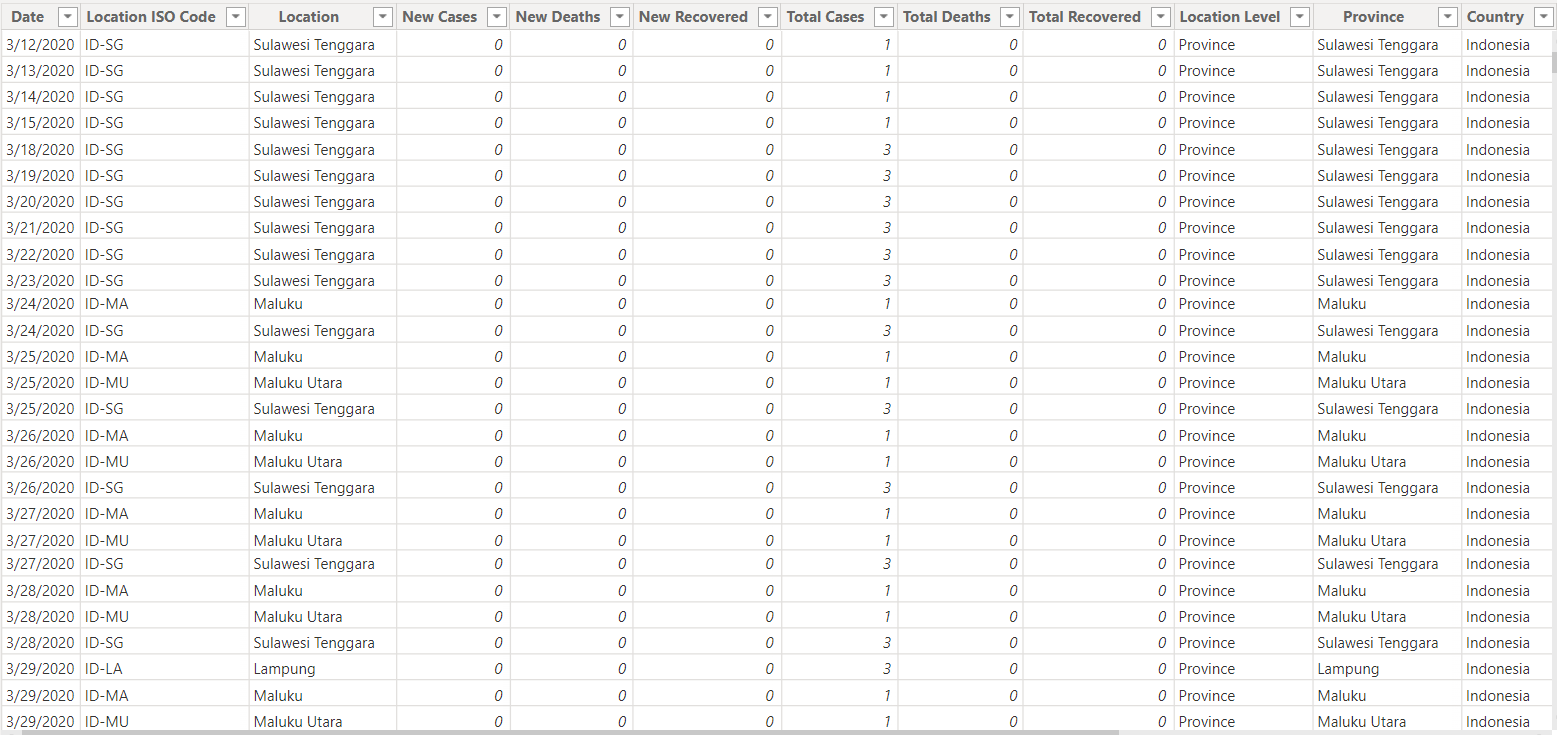
Penularan COVID-19 terjadi melalui droplet kecil yang keluar dari hidung atau mulut saat seseorang batuk atau bersin. Droplet ini dapat jatuh ke permukaan benda di sekitar kita. Jika orang lain menyentuh benda yang terkontaminasi oleh droplet tersebut dan kemudian menyentuh mata, hidung, atau mulut, maka orang tersebut dapat terinfeksi. Penularan juga bisa terjadi jika seseorang menghirup droplet dari orang yang terinfeksi. Oleh karena itu, menjaga jarak sosial sekitar satu meter dari orang yang mungkin terinfeksi sangat penting.

Saat ini, WHO menyatakan bahwa risiko penularan dari orang yang tidak menunjukkan gejala COVID-19 sangat kecil, meskipun ada beberapa kasus di mana orang hanya mengalami gejala ringan atau bahkan tidak merasa sakit sama sekali, terutama pada awal infeksi. Periode penularan atau masa inkubasi COVID-19 masih menjadi objek penelitian aktif oleh para ilmuwan. Untuk memperoleh informasi yang akurat tentang perkembangan penyakit ini, selalu penting untuk mengikuti sumber informasi resmi dan terpercaya.

Pada pembahasan selanjutnya yaitu menganalisis data penduduk Indonesia yang terjangkit virus Covid 19. Hal - hal yang harus dilakukan untuk menganalisis sebagai berikut :

1. Mencari Data

Data yang akan digunakan merupakan data yang berasal dari website Bernama Kaggle. Berikut data yang telah di ambil dari website tersebut.



Pada tabel diatas merupakan data pasien covid 19 yang berada di seluruh Indonesia. Data tersebut telah melalui beberapa proses yang telah dilakukan sebelumnya, sebelum data tersbut di ambil. Hal yang telah menjadikan data tersebut sangat penting adalah pembersihan data *(clean data)*. Hal ini dilakukan untuk menyesuaikan target yang akan dipenuhi sebelumnya untuk melihat persebaran covid 19 di Indonesia.

1. Clean Data

Clean data adalah data yang telah diperiksa, diproses, dan disesuaikan sedemikian rupa sehingga bebas dari kesalahan, inkonsistensi, atau duplikasi. Data yang bersih adalah data yang dapat diandalkan dan akurat, sehingga dapat digunakan untuk analisis, pelaporan, atau pengambilan keputusan tanpa risiko kesalahan yang disebabkan oleh data yang tidak terawat dengan baik.

Clean data sangat penting dalam berbagai konteks, termasuk analisis bisnis, penelitian ilmiah, dan pengambilan keputusan. Data yang tidak bersih atau tidak terawat dengan baik dapat menghasilkan analisis yang salah, pengambilan keputusan yang buruk, atau masalah lainnya dalam pemrosesan dan manajemen data. Oleh karena itu, pembersihan data merupakan langkah awal yang penting dalam pengelolaan data yang efektif.

1. Cara Mengatasi *Missing Value*

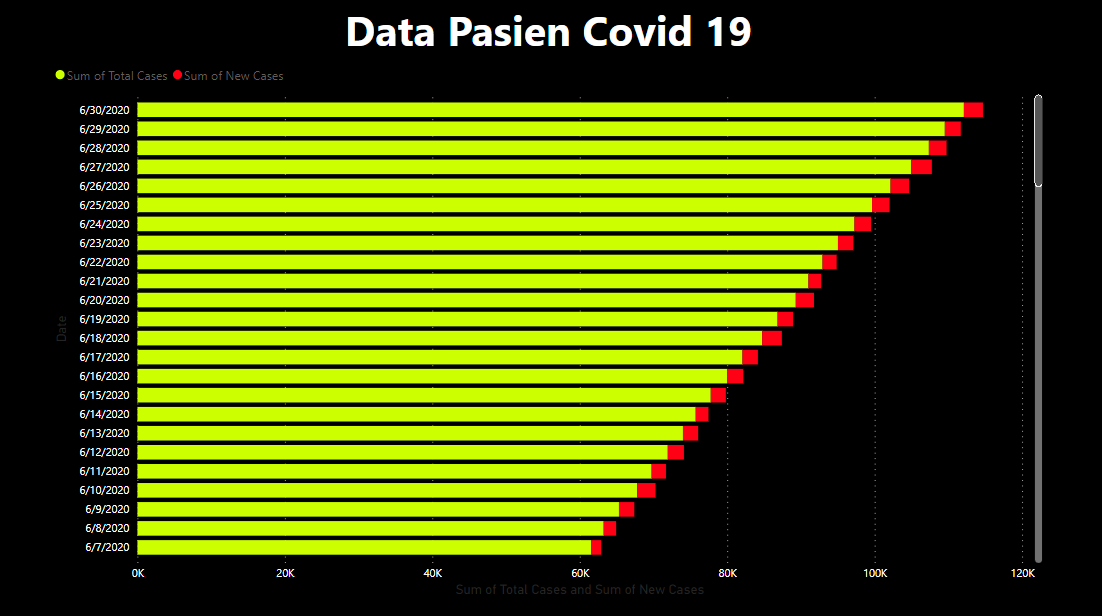
*Missing value* adalah nilai yang hilang atau tidak ada dalam data. Hal ini dapat terjadi karena berbagai alasan seperti kesalahan input, data yang tidak lengkap, atau data yang hilang karena alasan lain. Ketika ada *missing value* dalam data, dapat mempengaruhi hasil analisis data dan membuat kesimpulan yang salah. Ada beberapa cara untuk memperbaiki *missing value* dalam data, seperti yang kamu sebutkan tadi. Berikut adalah penjelasan lebih detail untuk masing-masing cara:

* + - * Mengisi dengan nilai 0 atau 1: Salah satu metode yang cukup sederhana untuk menangani data yang hilang adalah dengan mengganti nilai yang kosong dengan angka 0 atau 1. Tetapi, perlu diperhatikan bahwa metode ini lebih efektif jika data yang hilang terkait dengan variabel yang memiliki dua pilihan nilai, seperti variabel biner atau kategorikal.
      * Mengisi dengan nilai rata-rata atau *median*: Salah satu metode alternatif untuk menangani data yang hilang adalah dengan menggantinya menggunakan nilai rata-rata atau nilai tengah (median) dari variabel yang bersangkutan. Ini biasanya diterapkan ketika data yang hilang berhubungan dengan variabel numerik. Tetapi, penting untuk diingat bahwa pendekatan ini bisa memiliki dampak pada hasil analisis data.
      * Menggunakan *Machine Learning*: Salah satu pendekatan yang lebih advanced untuk menangani missing value adalah dengan memanfaatkan teknik-teknik Machine Learning seperti Regresi atau K-Means Clustering. Pendekatan ini dapat digunakan untuk memprediksi nilai yang hilang berdasarkan informasi yang sudah ada dalam dataset. Namun, metode ini melibatkan proses yang lebih kompleks dalam pemrosesan data dan memerlukan pemahaman yang lebih mendalam tentang konsep-konsep Machine Learning.

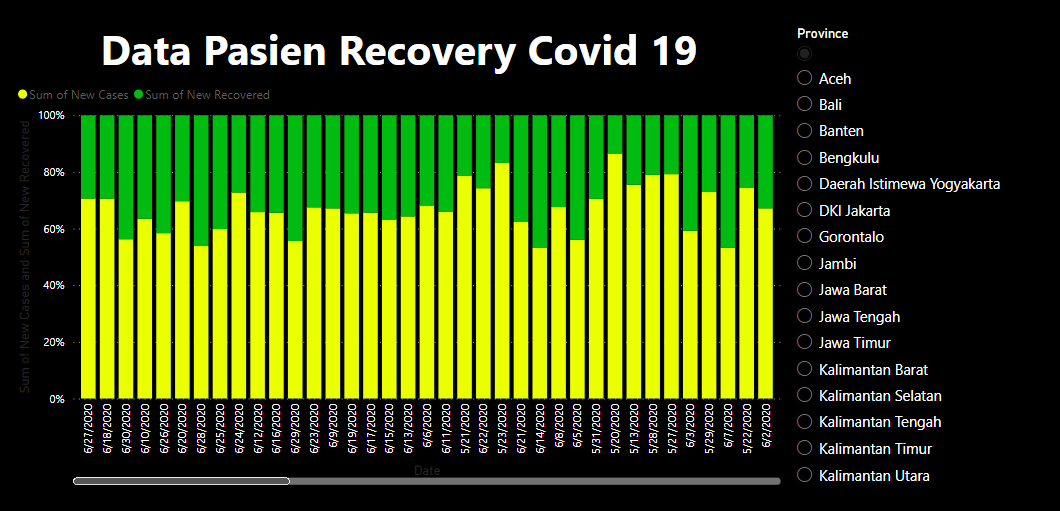
1. *Balancing* Data

Balancing data adalah proses untuk mengatasi ketidakseimbangan dalam distribusi kelas atau label pada dataset. Ketidakseimbangan ini terjadi ketika satu kelas atau label memiliki jumlah sampel yang jauh lebih banyak atau lebih sedikit dibandingkan dengan kelas atau label lainnya dalam dataset. Masalah ini sering terjadi dalam konteks klasifikasi, di mana kelas minoritas (kelas yang memiliki jumlah sampel lebih sedikit) cenderung memiliki dampak yang kurang signifikan dalam model pembelajaran mesin. Namun pada kasus kali ini tidak ada terjadi balancing data.

1. Analisis Data
2. Jumlah Kasus Covid 19 di Indonesia

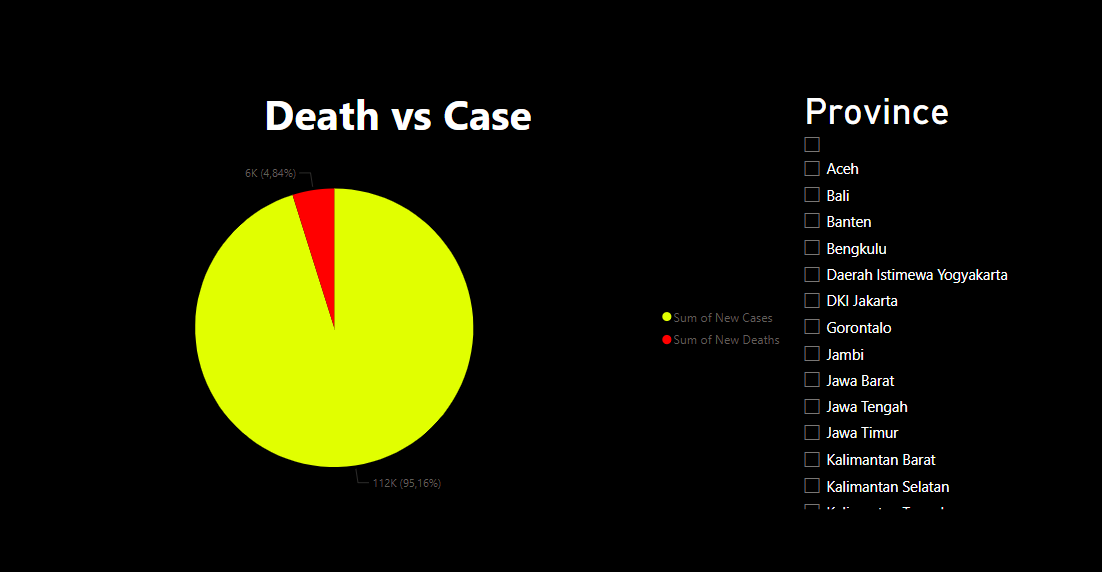
Gambar di atas adalah grafik yang menunjukkan jumlah penduduk Indonesia yang terinfeksi COVID-19 selama tahun 2020. Dari tanggal yang tertera di sebelahnya, kita dapat melihat bahwa mulai dari periode tertentu, angka kasus COVID-19 di Indonesia mengalami peningkatan yang sangat signifikan, menyebabkan lonjakan jumlah penduduk yang terpapar oleh virus ini. Hasil analisis pada grafik ini menampilkan dua jenis data yang berbeda. Yang pertama, dalam warna kuning, adalah total kasus harian penduduk yang terpapar COVID-19. Sedangkan yang kedua, yang ditunjukkan dalam warna merah, adalah jumlah kasus terbaru dari penduduk yang terinfeksi virus COVID-19. Tren peningkatan jumlah kasus COVID-19 yang begitu tajam dalam grafik ini menggarisbawahi signifikansi masalah ini bagi Indonesia. Analisis data semacam ini sangat penting untuk memahami perkembangan pandemi dan untuk mengambil langkah-langkah yang tepat dalam penanganannya.

1. Jumlah Sembuh vs Kasus Baru



Gambar di atas adalah analisis perkembangan kasus COVID-19, khususnya fokus pada jumlah pasien yang telah sembuh dari virus tersebut. Diagram tersebut menggambarkan peningkatan kasus baru dengan kecepatan yang sangat tinggi, bahkan melebihi jumlah pasien yang telah pulih. Pada diagram, area berwarna kuning mewakili jumlah kasus baru harian, yang dianalisis berdasarkan tanggal atau setiap harinya. Di sisi lain, area berwarna hijau menggambarkan jumlah orang yang telah sembuh dari COVID-19. Namun, perlu diperhatikan bahwa jumlah pasien yang sembuh dari COVID-19, yang ditunjukkan oleh area hijau, masih belum mencapai target yang diharapkan oleh pemerintah Indonesia. Ini menunjukkan bahwa upaya untuk mengatasi pandemi masih perlu ditingkatkan, seperti peningkatan vaksinasi, penerapan protokol kesehatan yang ketat, dan peran aktif masyarakat dalam memutus rantai penularan virus. Data visualisasi seperti ini sangat penting untuk memahami dinamika pandemi dan membantu pengambilan keputusan pemerintah dalam menangani situasi yang terus berkembang. Perlu kerjasama dari berbagai pihak untuk mencapai target pengendalian COVID-19 yang diinginkan.

1. Jumlah Kematian vs Kasus Baru



Pada grafik di atas, kita dapat melihat perbandingan antara jumlah kasus kematian akibat COVID-19 di seluruh Indonesia. Menariknya, jumlah kasus kematian tampaknya masih relatif rendah dibandingkan dengan jumlah kasus baru yang terpapar virus ini. Dalam grafik lingkaran, kita bisa membedakan perbandingan ini dengan warna kuning yang mewakili kasus terbaru, yang mengindikasikan orang-orang yang baru-baru ini terinfeksi COVID-19. Sementara itu, bagian yang berwarna merah menggambarkan total kasus kematian akibat COVID-19 secara keseluruhan di Indonesia. Ini menunjukkan bahwa meskipun ada peningkatan kasus baru, tingkat kematian masih tetap terkendali. Namun, perlu dicatat bahwa grafik ini hanya merupakan gambaran kasar, dan kita perlu melihat lebih dalam ke dalam data dan faktor-faktor lain yang mungkin memengaruhi tingkat kematian. Tetap penting untuk terus mengikuti perkembangan situasi dan langkah-langkah pencegahan yang dianjurkan oleh otoritas kesehatan untuk mengendalikan penyebaran virus ini.