



Pengantar Statistika - Menyajikan Data

Kelompok 4, Prodi Manajemen Informatika '24, Ilmu Komputer, Universitas Lampung



Kelompok 4 :


- Amrullah Azzaky (2477051027)
- Qorina Qisthi Kamal (2477051028)
- Arif Ahmad Muzakky (2477051017)
- AlbuQori Muhammad Anas Saragih (2477051026)
- Defina Rahmayanti (2477051023)

Pengenalan Statistika dan Menyajikan data



Sejarah Statistika

- **Era Kuno (3000 SM - 500 M), di Babilonia dan Mesir** -> Digunakan untuk mencatat data jumlah penduduk, hasil panen, dan pajak untuk kepentingan pemerintahan. Statistiknya berbentuk daftar atau catatan berurutan.
- **Periode Abad Pertengahan (1600 M - 1700 M), di Eropa** -> Digunakan untuk mencatat sensus penduduk, data ekonomi, dan administrasi pemerintahan. Statistiknya berbentuk daftar numerik dan mulai dianalisis secara lebih sistematis, seperti perhitungan angka kelahiran dan kematian.
- **Periode Statistik Klasik (1700 M - 1800 M), di Eropa** -> Digunakan untuk analisis ekonomi, sosial, dan pemerintahan. Statistiknya berbentuk tabel numerik dan mulai dikembangkan metode probabilitas serta distribusi data oleh Jakob Bernoulli dan Pierre-Simon Laplace.

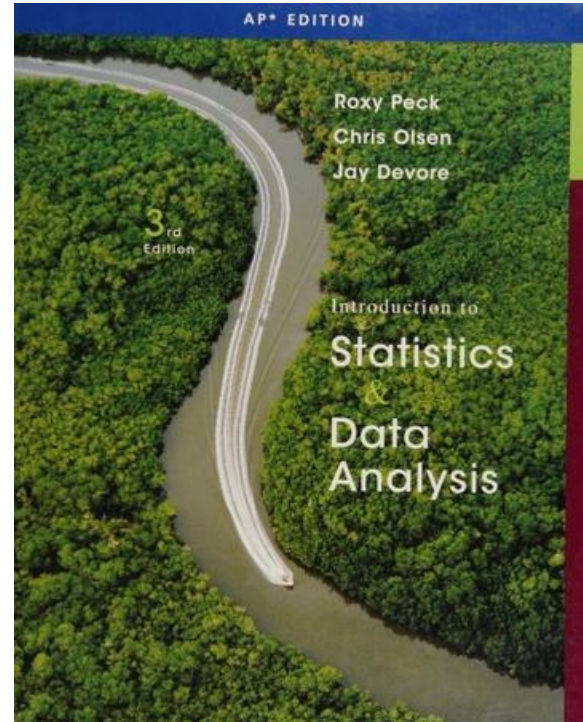
- 
- **Perkembangan Modern (1800 M - Sekarang), di Seluruh Dunia** -> Digunakan dalam berbagai bidang seperti sains, ekonomi, bisnis, dan teknologi. Statistiknya berbentuk tabel, grafik, dan model matematis kompleks. Metode seperti korelasi, regresi, dan inferensial diperkenalkan oleh Karl Pearson dan Ronald A. Fisher, serta semakin berkembang dengan bantuan komputer dan kecerdasan buatan.

Pengertian Statistika

Dalam buku *“Introduction to statistic and Data Analysis”* Statistika adalah disiplin ilmiah yang menyediakan metode, untuk, kita memahami sebuah atau banyak data.

Terdapat dua cabang utama statistika dalam buku ini, yaitu :

- Statistika Deskriptif
- Statistika Inferensial





Cabang Utama Statistika

Statistika Deskriptif

Statistika deskriptif adalah metode untuk mengumpulkan, mengklasifikasikan, meringkas, menginterpretasikan, dan menyajikan data. Tujuannya adalah untuk memberikan gambaran umum tentang data yang tersedia.

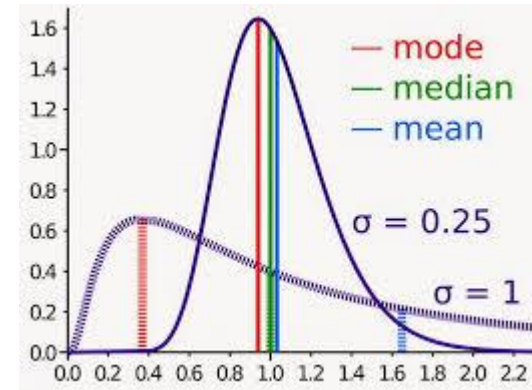
Statistika Inferensial

Statistika inferensial adalah metode untuk menarik kesimpulan tentang populasi dengan menggunakan data sampel. Metode ini menggunakan teori probabilitas dan model statistik.

Cabang Utama Statistika



Statistika Deskriptif



Statistika Inferensial



Cabang Utama Statistika

Statistika Deskriptif -> Hanya menggambarkan data yang ada tanpa menarik kesimpulan lebih luas.

Contoh: Rata-rata nilai ujian mahasiswa di kelas adalah 75.

Teknik: Mean (rata-rata), median, modus, diagram batang, histogram.

Statistika Inferensial -> Menggunakan data sampel untuk membuat kesimpulan tentang populasi yang lebih besar.

Contoh: Berdasarkan survei 100 orang, 60% masyarakat menyukai produk baru.

Teknik: Uji hipotesis, estimasi, regresi, probabilitas.



Menyajikan Data Secara Visual

Visualisasi data adalah teknik untuk menyajikan data dalam bentuk grafik, diagram, atau plot sehingga lebih mudah dipahami dan dianalisis.



Menyajikan Data Secara Visual

Pentingnya Penyajian Data secara Visual ?

- Memudahkan dalam melihat pola dan tren dalam data.
- Membantu dalam membandingkan kategori atau nilai tertentu.
- Mempermudah pengambilan keputusan berdasarkan data. Terkhusus jika sebuah perusahaan dan perusahaan tersebut memiliki atasan



Menyajikan Data Secara Visual

Diagram Batang (Bar Chart)

Digunakan untuk menyajikan data kategori. Setiap kategori direpresentasikan dengan batang yang panjangnya sesuai dengan jumlah atau frekuensi data tersebut.

Contoh: Jumlah mahasiswa berdasarkan jurusan di suatu universitas.

Diagram Lingkaran (Pie Chart)

Menampilkan proporsi setiap kategori dalam bentuk lingkaran yang dibagi menjadi beberapa sektor. Biasanya digunakan untuk menunjukkan persentase suatu kategori dari keseluruhan data.

Contoh: Persentase pengguna berbagai merek ponsel di suatu daerah.

Histogram

Digunakan untuk data numerik dan menampilkan distribusi data. Mirip dengan diagram batang, tetapi batangnya saling berdekatan tanpa celah karena data bersifat kontinu.

Contoh: Distribusi tinggi badan mahasiswa dalam suatu kelas.

Menyajikan Data Secara Visual

Diagram Batang (Bar Chart)

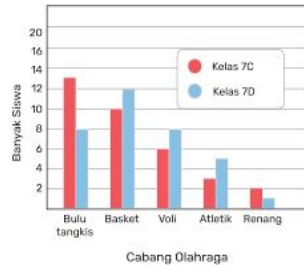
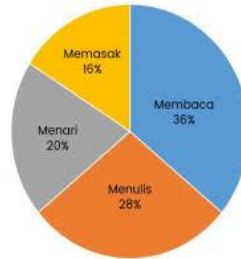
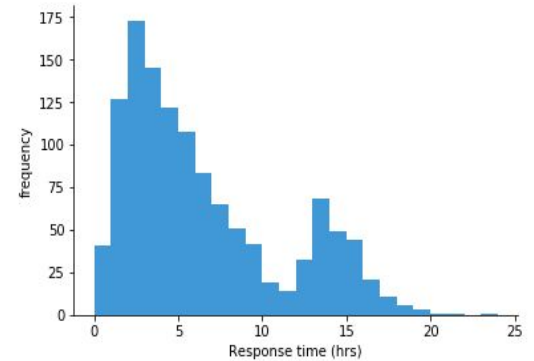


Diagram Lingkaran (Pie Chart)



Histogram



Sumber Penyajian Data, Kelompok 4 - WHO - COVID 19 Global Daily Data



WHO - COVID 19 Global Daily Data

Data merupakan sekumpulan fakta, angka, teks, atau simbol yang belum diproses dan belum memiliki makna tertentu. Data bisa berupa angka dalam spreadsheet, teks dalam dokumen, atau bahkan gambar dan suara. Setelah diproses dan dianalisis, data bisa menjadi informasi yang berguna untuk pengambilan keputusan.

Organisasi kesehatan dunia, *WHO*, menyediakan data statistik *Covid-19* yang dapat kita unduh, bersifat terbuka; umum; *open source*; gratis. Pada saat wabah covid melanda, *WHO* terus memperbaharui data tersebut secara berkala. Data ini memuat jumlah kasus, kematian, dan data vaksinasi *Covid-19*.

sumber data: <https://data.who.int/dashboards/covid19/data>



COVID 19 Global Daily Data - CSV Format

- **WHO** menyediakan data statistik Covid-19 yang dapat kita unduh berformatkan CSV.
- **CSV** atau **Comma Separated Value** sendiri adalah format file yang menyimpan data dalam bentuk teks biasa, dengan setiap nilai yang dipisahkan oleh koma (",") atau titik ("."). Kelebihannya:
 - > Mudah memindahkan data dari satu program ke program lain
 - > Dapat dibuka dengan berbagai penyunting teks seperti Notepad, Wordpad, ataupun Microsoft Excel
 - > Kompatibel dengan semua bahasa pemrograman
 - > Fleksibel sehingga dapat dimanipulasi dan dianalisis dengan mudah

```
File Edit View
ID, Nama, Umur, Kota, Pekerjaan, Gaji
1, Andi, 25, Jakarta, Software Engineer, 12000000
2, Budi, 30, Bandung, Data Analyst, 10000000
3, Citra, 28, Surabaya, Digital Marketer, 9000000
4, Dewi, 27, Yogyakarta, UI/UX Designer, 9500000
5, Eka, 35, Medan, Project Manager, 15000000
6, Fajar, 29, Semarang, System Administrator, 11000000
7, Gina, 26, Bali, Content Writer, 8500000
8, Hadi, 31, Malang, Network Engineer, 10500000
9, Indah, 24, Bogor, Social Media Specialist, 8000000
10, Joko, 33, Palembang, Cybersecurity Analyst, 13000000
```

Visualisasi Data, Statistik Deskriptif - COVID 19 Global Daily Data



matplotlib

Pandas





seaborn



Tools Visualisasi Data

- **Python**, bahasa pemrograman tingkat tinggi, dengan sintaksnya yang sederhana.
- **Matplotlib**, merupakan pustaka (library) dalam Python yang digunakan untuk membuat berbagai jenis visualisasi data. Digunakan untuk membuat histogram.
- **Pandas**, adalah pustaka (library) Python yang digunakan untuk manipulasi dan analisis data. Digunakan untuk membaca *data set* dari file CSV, Memfilter data untuk Indonesia, mengubah format tanggal, melakukan analisis statistik deskriptif, mengelompokkan dan membandingkan data antar negara.



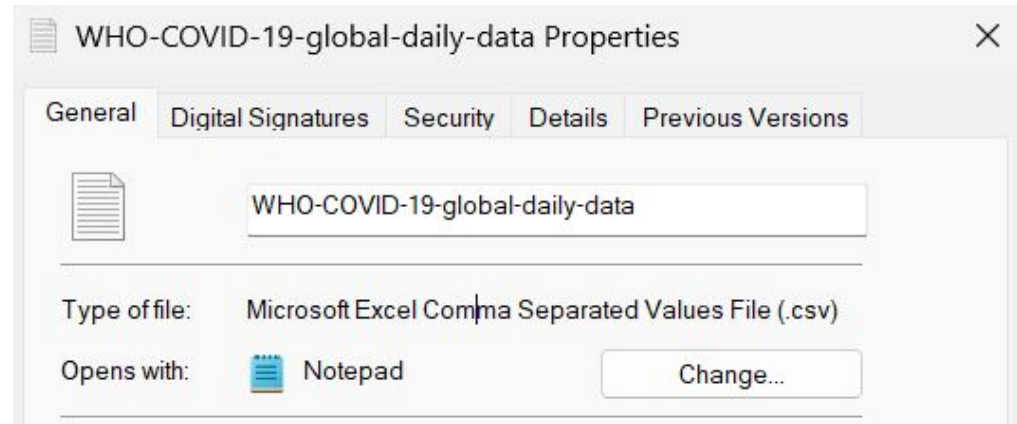
Tools Visualisasi Data

- **Seaborn** merupakan library Python berbasis Matplotlib yang digunakan untuk membuat visualisasi data yang informatif.
- **Jupyter Notebook** adalah IDE yang sering digunakan untuk analisis data, machine learning, dan pengolahan data dengan Python.

Data Set

Data set merupakan kumpulan data yang terstruktur dan berisi informasi dari berbagai sumber. Data set dapat berupa data numerik, teks, gambar, atau gabungan dari semuanya.

Dalam konteks kali ini, **data set** merujuk pada kumpulan teks dan angka, informasi *Covid-19*.



File Edit Selection View Go Run ... ProjectPythonPengantarStatistikaKelompok

EXPLORER ... analisis_covid19.ipynb analisisCovid.ipynb WHO-COVID-19-global-daily-data.csv Extension: Rainbow CSV Waktu Sholat Maghrib Telah Tiba

PROJECTPYTHONPENGAN... WHO-COVID-19-global-daily-data.csv > data

.ipynb_checkpoints
analisisCovid.ipynb
outputKasus.png
outputKasus2.png
outputKasus3.png
WHO-COVID-19-...

```
1 Date_reported, Country_code, Country, WHO_region, New_cases, Cumulative_cases, New_deaths, Cumulative_deaths
2 2020-01-04, AI, Anguilla, AMR, 0, 0
3 2020-01-04, AZ, Azerbaijan, EUR, 0, 0
4 2020-01-04, BD, Bangladesh, SEAR, 0, 0, 0
5 2020-01-04, BB, Barbados, AMR, 0, 0
6 2020-01-04, BR, Brazil, AMR, 0, 0
7 2020-01-04, CA, Canada, AMR, 0, 0
8 2020-01-04, CO, Colombia, AMR, 0, 0
9 2020-01-04, KM, Comoros, AFR, 0, 0
10 2020-01-04, CK, Cook Islands, WPR, 0, 0, 0
11 2020-01-04, CW, Curaçao, AMR, 0, 0
12 2020-01-04, DM, Dominica, AMR, 0, 0
13 2020-01-04, DO, Dominican Republic, AMR, 0, 0
14 2020-01-04, FK, Falkland Islands, AMR, 0, 0
15 2020-01-04, GA, Gabon, AFR, 0, 0
16 2020-01-04, GP, Guadeloupe, AMR, 0, 0
17 2020-01-04, IR, Iran (Islamic Republic of), EMR, 0, 0
18 2020-01-04, IE, Ireland, EUR, 0, 0
19 2020-01-04, IM, Isle of Man, EUR, 0, 0
20 2020-01-04, IN, India, SEAR, 0, 0, 0
21 2020-01-04, JO, Jordan, EMR, 0, 0
22 2020-01-04, LS, Lesotho, AFR, 0, 0
23 2020-01-04, MW, Malawi, AFR, 0, 0
24 2020-01-04, QA, Qatar, EMR, 0, 0
25 2020-01-04, NR, Nauru, WPR, 0, 0, 0
26 2020-01-04, RO, Romania, EUR, 0, 0
27 2020-01-04, RW, Rwanda, AFR, 0, 0
28 2020-01-04, VC, Saint Vincent and the Grenadines, AMR, 0, 0
29 2020-01-04, SN, Senegal, AFR, 0, 0
30 2020-01-04, SB, Solomon Islands, WPR, 0, 0, 0
31 2020-01-04, LK, Sri Lanka, SEAR, 0, 0, 0
32 2020-01-04, SY, Syrian Arab Republic, EMR, 0, 0
33 2020-01-04, TJ, Tajikistan, EUR, 0, 0
34 2020-01-04, TH, Thailand, SEAR, 0, 0, 0
35 2020-01-04, AE, United Arab Emirates, EMR, 0, 0
36 2020-01-04, TZ, United Republic of Tanzania, AFR, 0, 0
37 2020-01-04, VE, Venezuela (Bolivarian Republic of), AMR, 0, 0
38 2020-01-04, AT, Austria, EUR, 0, 0
39 2020-01-04, BE, Belgium, EUR, 0, 0
40 2020-01-04, BM, Bermuda, AMR, 0, 0
41 2020-01-04, BF, Burkina Faso, AFR, 0, 0
42 2020-01-04, BI, Burundi, AFR, 0, 0
43 2020-01-04, KH, Cambodia, WPR, 0, 0, 0
44 2020-01-04, TD, Chad, AFR, 0, 0
45 2020-01-04, DK, Denmark, EUR, 0, 0
46 2020-01-04, GF, French Guiana, AMR, 0, 0
47 2020-01-04, GD, Grenada, AMR, 0, 0
48 2020-01-04, JP, Japan, WPR, 0, 0, 0
49 2020-01-04, JE, "Jersey, Channel Islands", EUR, 0, 0
```

> OUTLINE
> TIMELINE
> VS CODE PETS

Col 1: Date_reported Ln 1, Col 1 Spaces: 4 UTF-8 with BOM CRLF {} CSV Maghrib(18:11)-00:00:00 [KOTA JAKARTA] Prettier

Data Set, COVID 19 Global Daily Data - CSV



Data Set

Dataset ini, *COVID 19 Global Daily Data*, memiliki 449.040 baris dan 8 kolom. Dengan rincian per kolom sebagai berikut: (Kolom A-H)

1. Date_reported: Tanggal pelaporan data.

2. Country_code: Kode negara.

3. Country: Nama negara.

4. WHO_region: Wilayah WHO.

5. New_cases: Kasus baru harian.

6. Cumulative_cases: Total kasus kumulatif.

7. New_deaths: Kematian baru harian.

8. Cumulative_deaths: Total kematian kumulatif.



Istilah-Istilah yang akan digunakan

- **count:** Jumlah data
- **mean:** Rata-rata
- **std:** Standar deviasi
- **min, max:** Nilai minimum dan maksimum
- **25%, 50%, 75%:** Kuartil data
- **Kuartil:** Pembagi data
- **Median:** Nilai tengah

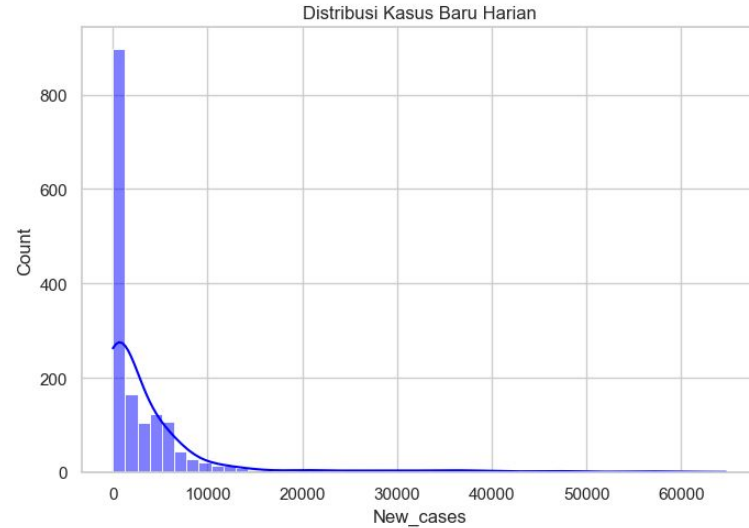
Visualisasi Data - Histogram

Kasus Harian Baru (New_cases)

Rata-rata: 4.177 kasus/hari

Median: 803 kasus/hari (menunjukkan distribusi yang miring ke kanan)

Max: 64.718 kasus



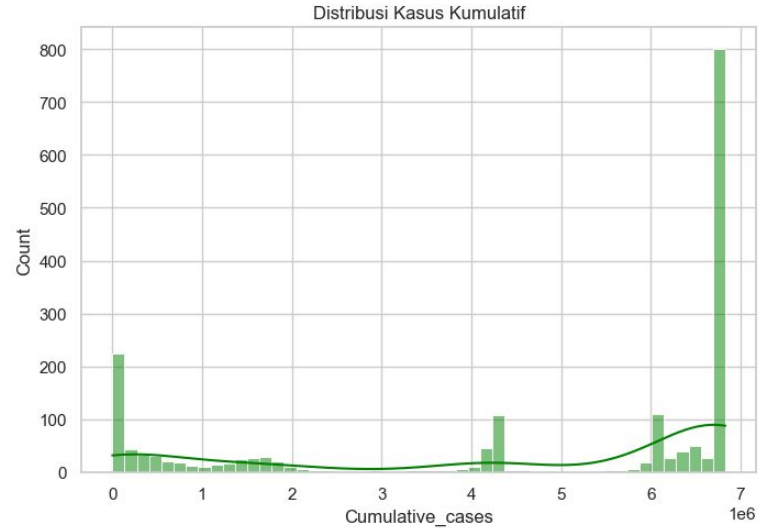
Visualisasi Data - Histogram

Kasus Kumulatif (Cumulative_cases)

Rata-rata: 4,56 juta kasus

Median: 6,18 juta kasus

Max: 6,83 juta kasus (mendekati total kasus yang pernah tercatat)



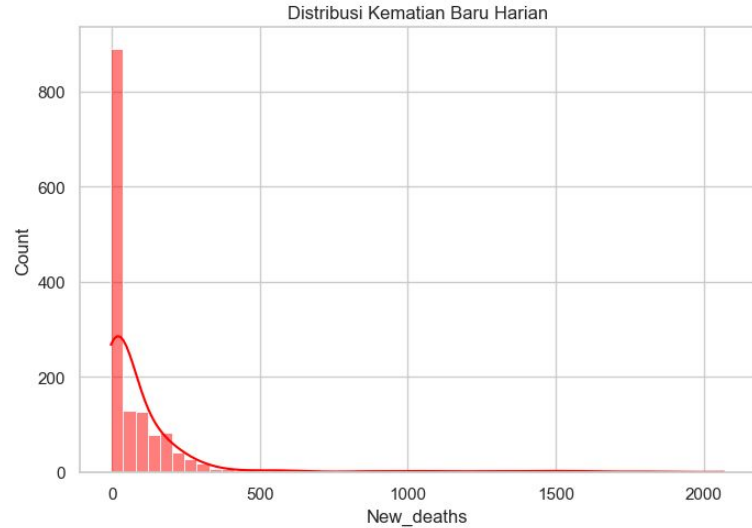
Visualisasi Data - Histogram

Kematian Harian Baru (New_deaths)

Rata-rata: 109 kematian/hari

Median: 19,5 kematian/hari

Max: 2.069 kematian/hari (pada puncak pandemi)



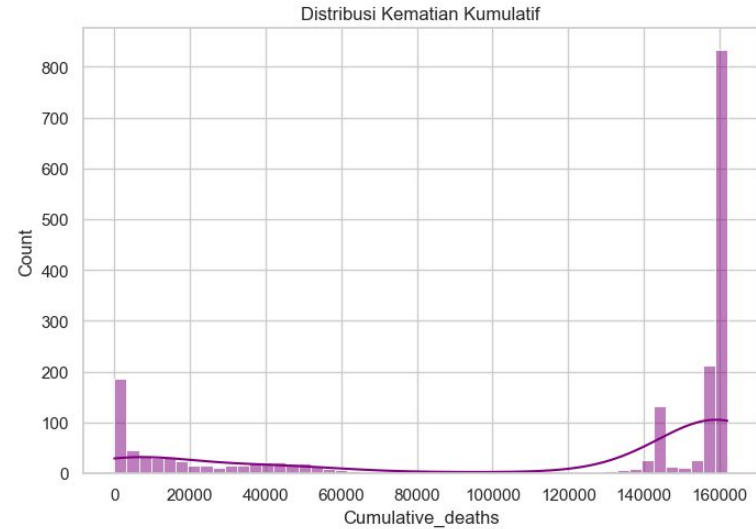
Visualisasi Data - Histogram

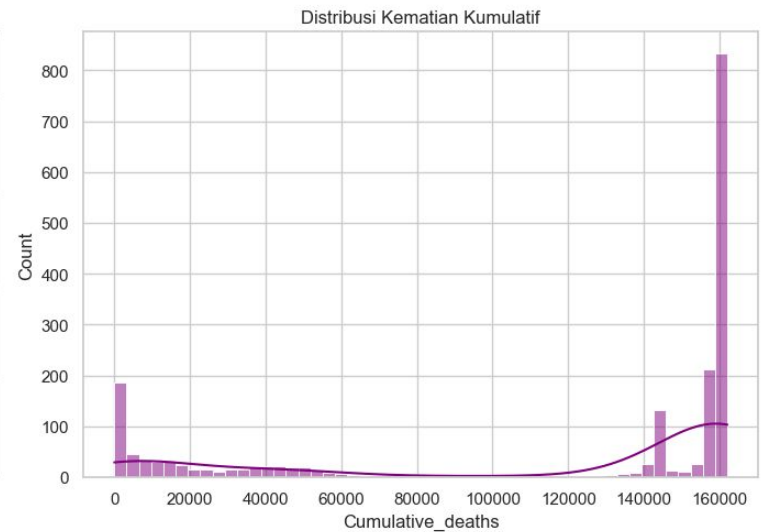
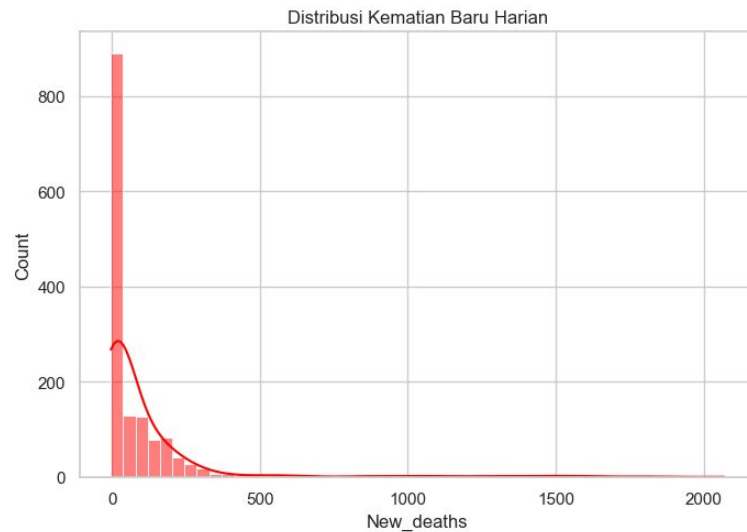
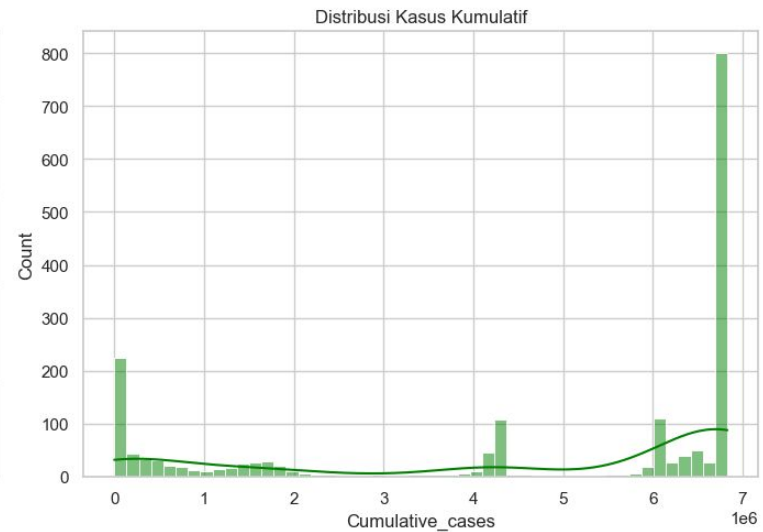
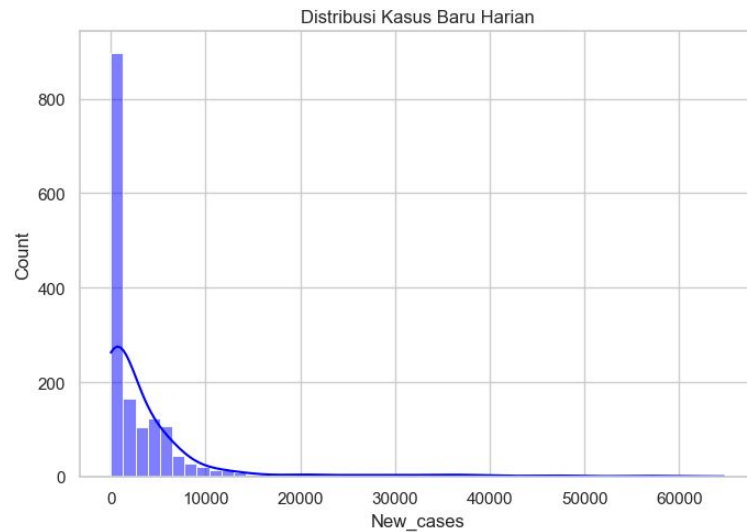
Kematian Kumulatif (Cumulative_deaths)

Rata-rata: 115.071 kematian

Median: 156.940 kematian

Max: 162.059 kematian







Membaca Pola Data

Kasus Baru Harian (New Cases)

> Distribusinya miring ke kanan, dengan beberapa lonjakan besar pada hari-hari tertentu.

Kematian Baru Harian (New Deaths)

> Polanya mirip dengan kasus baru, menunjukkan beberapa lonjakan besar pada waktu tertentu.

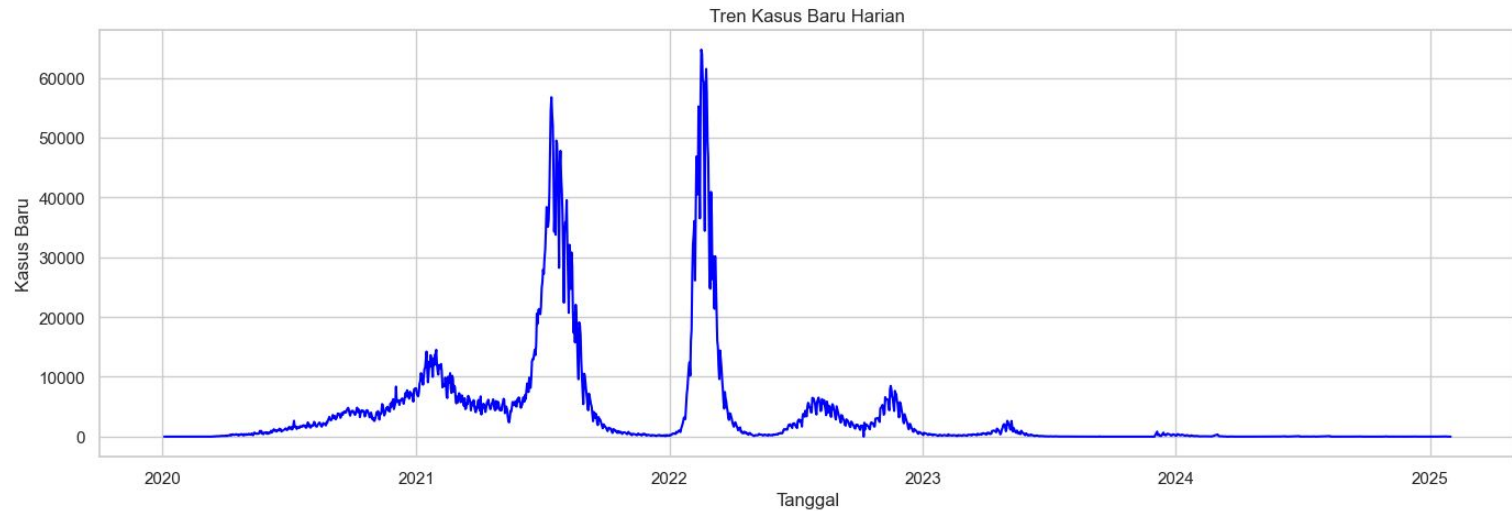
Kasus Kumulatif (Cumulative Cases)

> Distribusi terlihat naik secara eksponensial, yang masuk akal karena kasus kumulatif terus bertambah seiring waktu.

Kematian Kumulatif (Cumulative Deaths)

> Distribusinya stabil dan meningkat, dengan pola yang serupa dengan kasus kumulatif.

Visualisasi Data - Grafik Tren



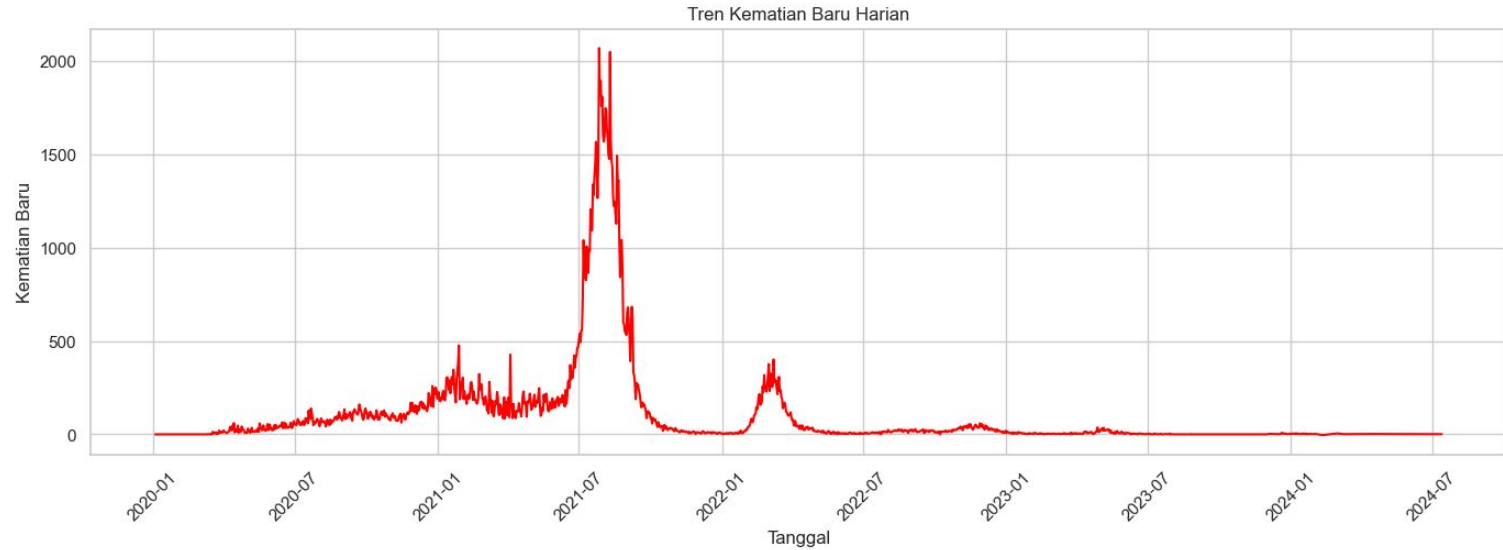


Visualisasi Data - Grafik Tren

Tren Kasus Baru Harian

- Terlihat beberapa lonjakan signifikan, kemungkinan saat gelombang besar COVID-19 terjadi (seperti varian Delta atau Omicron).
- Setelah puncak kasus, jumlah kasus harian cenderung menurun.

Visualisasi Data - Grafik Tren

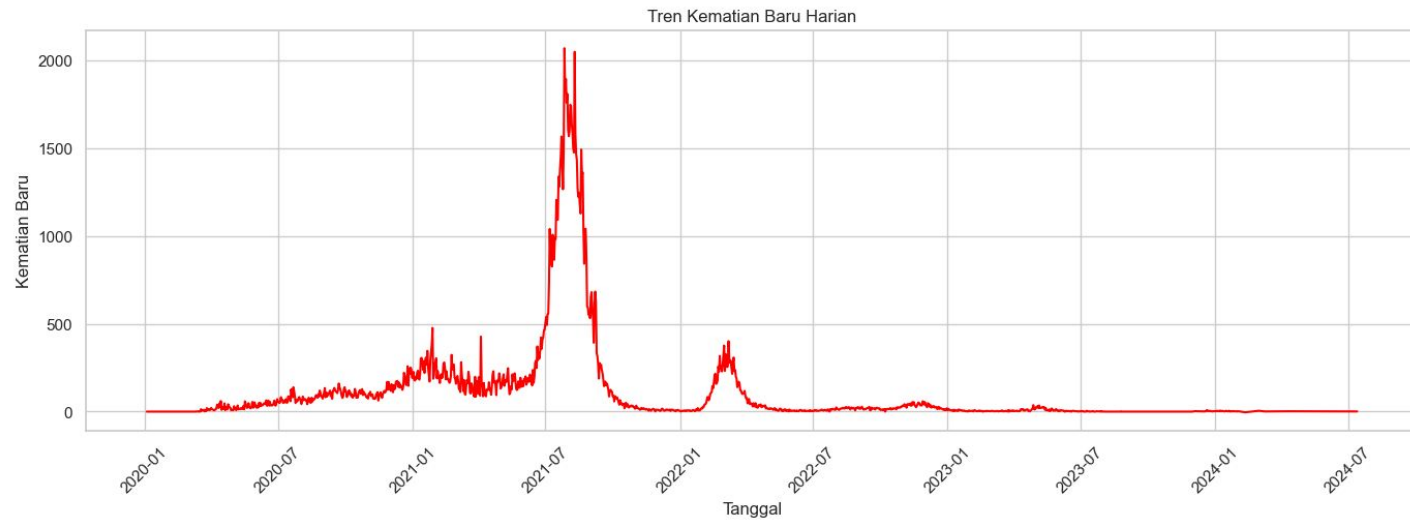
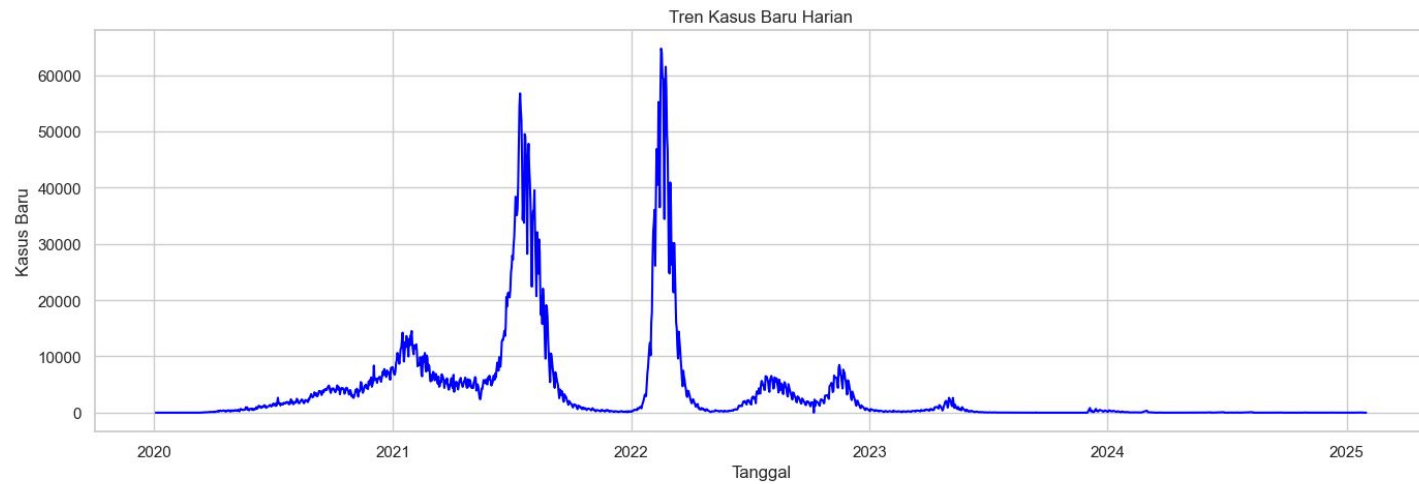




Visualisasi Data - Grafik Tren

Tren Kematian Baru Harian

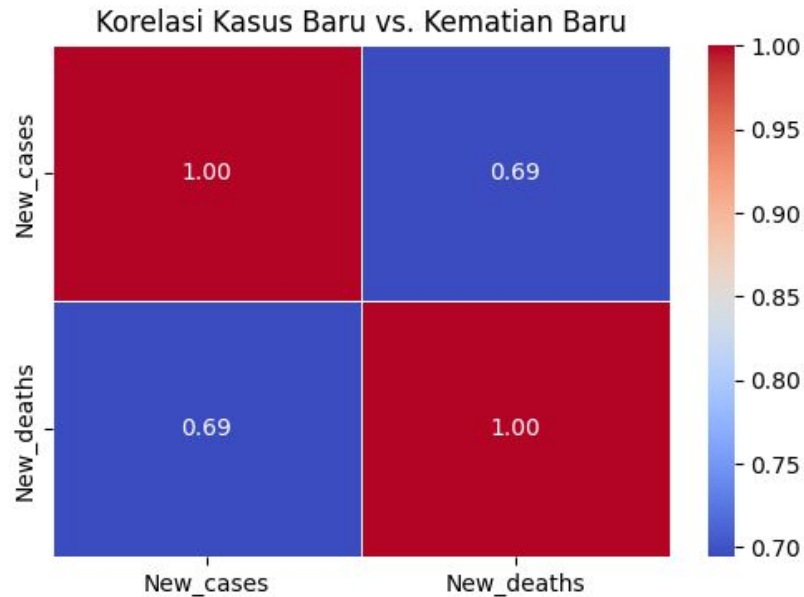
- Polanya mengikuti tren kasus baru yang masuk akal, logis, karena kematian biasanya terjadi setelah infeksi parah.
- Puncak kematian terbesar terjadi setelah lonjakan kasus tertinggi.



Korelasi Antara *New_Cases* dan *New_Deaths*

Hasil analisis korelasi menunjukkan bahwa kasus baru dan kematian baru memiliki korelasi sebesar 0,69. Ini berarti terdapat hubungan positif yang cukup kuat antara jumlah kasus harian dan jumlah kematian harian, semakin tinggi jumlah kasus baru, semakin tinggi pula jumlah kematian baru yang terjadi.

Namun, korelasi ini tidak mencapai 1, yang menunjukkan bahwa faktor lain juga berperan dalam jumlah kematian, seperti kapasitas sistem kesehatan, tingkat keparahan kasus, atau intervensi medis.



Revisi, Kelompok 4 - Kasus Harian Covid-19 Tertinggi

Kompas.com / News / Nasional

UPDATE 16 Februari: Rekor 64.718 Kasus Baru Covid-19, Tertinggi Selama Pandemi

Kompas.com - 16/02/2022, 18:18 WIB

 **Mutia Fauzia, Diamanty Meiliana**
Tim Redaksi





Rekor Tertinggi Kasus Covid-19 Dalam Satu Hari

Pada 16 Februari 2022, Rabu, Indonesia mencatat rekor tertinggi kasus harian COVID-19 dengan penambahan **64.718 kasus baru**, sehingga total kasus mencapai 4.966.046. Provinsi dengan penambahan kasus tertinggi adalah Jawa Barat (15.196 kasus), diikuti oleh DKI Jakarta (12.388 kasus) dan Jawa Timur (7.919 kasus).

Pada hari yang sama, terdapat 25.386 pasien sembuh, meningkatkan total kesembuhan menjadi 4.375.234. Namun, jumlah kasus aktif juga meningkat menjadi 445.190 kasus.

sumber: <https://nasional.kompas.com/read/2022/02/16/18180721/update-16-februari-rekor-64718-kasus-baru-covid-19-tertinggi-selama-pandemi>

```
stats = {
    "Tanggal": max_case_day["Date_reported"],
    "Negara": max_case_day["Country"],
    "Kasus Baru": int(max_case_day["New_cases"]),
    "Kasus Kumulatif": int(max_case_day["Cumulative_cases"]),
    "Kematian Baru": int(max_case_day["New_deaths"]),
    "Kematian Kumulatif": int(max_case_day["Cumulative_deaths"]),
}
```

```
df_stats = pd.DataFrame([stats])
df_stats
```

[9] ✓ 0.0s

...

	Tanggal	Negara	Kasus Baru	Kasus Kumulatif	Kematian Baru	Kematian Kumulatif
0	2022-02-16	Indonesia	64718	4966046	167	145622



Tabel Kasus Harian Covid-19 Tertinggi

Date_reported	Country	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
2022-02-16	Indonesia	64.718	4.966.046	167	145.622

- **Puncak Kasus:** 16 Februari 2022 mencatat jumlah kasus baru tertinggi sepanjang pandemi di Indonesia dengan 64.718 kasus.
- **Dampak Akumulatif:** Pada tanggal tersebut, jumlah kasus kumulatif telah mencapai 4.966.046 kasus, menunjukkan tingginya penyebaran COVID-19 sejak awal pandemi.
- **Kematian:** Sebanyak 167 kematian baru terjadi pada hari tersebut, dengan total kematian mencapai 145.622 jiwa.

File Edit Selection View Go Run ...

ProjectPythonPengantarStatistikaKelompok

ovid19.ipynb analisisCovid.ipynb Waktu Sholat Isya Telah Tiba WHO-COVID-19-global-daily-data.csv WHO-COVID-19-global-daily-data-checkpoint.csv

2022-02-15.ID

Replace AB

2 results in 2 files - Open in editor

> WHO-COVID-19-global-daily-data-checkpoint.csv

> WHO-COVID-19-global-daily-data-checkpoint.csv

2022-02-15.ID

.ipynb_checkpoints > WHO-COVID-19-global-daily-data-checkpoint.csv > data

	Date_reported	Country_code	Country	WHO_region	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
185642	2022-02-15	PH	Philippines	WPR	1990	3041920	32	55140
185643	2022-02-15	MP	Northern Mariana Islands	WPR	159	7746	1	25
185644	2022-02-15	MK	North Macedonia	EUR	287380	9198		
185645	2022-02-15	MM	Myanmar	SEAR	2467	550824	3	19314
185646	2022-02-15	MT	Malta	EUR	69955	610		
185647	2022-02-15	MV	Maldives	SEAR	1037	161113	0	289
185648	2022-02-15	LI	Liechtenstein	EUR	119	10890	73	
185649	2022-02-15	ID	Indonesia	SEAR	57049	4901328	134	145455
185650	2022-02-15	IT	Italy	EUR	28776	12134451	281	151296
185651	2022-02-15	GY	Guyana	AMR	34	62255	1	1194
185652	2022-02-15	HU	Hungary	EUR	4411	1721483	123	42754
185653	2022-02-15	GG	"Guernsey, Channel Islands"	EUR	110	12730	31	
185654	2022-02-15	GH	Ghana	AFR	158434	1438		
185655	2022-02-15	GI	Gibraltar	EUR	119	14464	101	
185656	2022-02-15	GM	Gambia	AFR	11924	365		
185657	2022-02-15	CZ	Czechia	EUR	23226	3611440	49	38150
185658	2022-02-15	CY	Cyprus	EUR	2930	291865	3	801
185659	2022-02-15	CV	Cabo Verde	AFR	4	55846	400	
185660	2022-02-15	BJ	Benin	AFR	26567	163		
185661	2022-02-15	BW	Botswana	AFR	1422	261913	5	2608
185662	2022-02-15	AD	Andorra	EUR	137	37277	1	149
185663	2022-02-15	AR	Argentina	AMR	15935	8783109	76	127529
185664	2022-02-15	AF	Afghanistan	EMR	472	171076	7	7501
185665	2022-02-15	ZW	Zimbabwe	AFR	222	231603	5374	
185666	2022-02-15	VN	Viet Nam	WPR	31814	2572087	85	39122
185667	2022-02-15	TM	Turkmenistan	EUR	0	0	0	
185668	2022-02-15	TO	Tonga	WPR	0	141	0	0
185669	2022-02-15	WS	Samoa	WPR	0	33	0	0
185670	2022-02-15	ST	Sao Tome and Principe	AFR	5920	71		
185671	2022-02-15	SO	Somalia	EMR	26203	1340		
185672	2022-02-15	SH	Saint Helena	AFR	0	0		
185673	2022-02-15	BL	Saint Barthelemy	AMR	3633	4		

Ln 185649, Col 54 (6 selected) Spaces: 4 UTF-8 with BOM CRLF {} CSV Isya(19:19): 00:00:00 [KOTA JAKARTA] Prettier

Kasus Harian Covid-19, Satu Hari Sebelum di Hari Kasus Tertinggi - CSV



Tabel Kasus Harian Covid-19

Date_reported	Country	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
2022-02-16	Indonesia	64.718	4.966.046	167	145.622
2022-02-15	Indonesia	57.049	4.901.328	134	145.455

Karena satu data di suatu hari, **16-02-2022**, terlalu sempit untuk kita ambil kesimpulan numeriknya. Kita akan menambahkan satu data lagi di satu hari sebelumnya, **15-02-2022**.



Persamaan Numerik

$$\text{Persentase Kenaikan (\%)} = \frac{(\text{Akhir} - \text{Awal})}{\text{Awal}} \times 100\%$$



Persamaan Numerik

$$\text{Persentase Kenaikan} = \frac{\text{Kasus Hari Tertinggi} - \text{Kasus Hari Sebelumnya}}{\text{Kasus Hari Sebelumnya}} \times 100\%$$

Persamaan Numerik

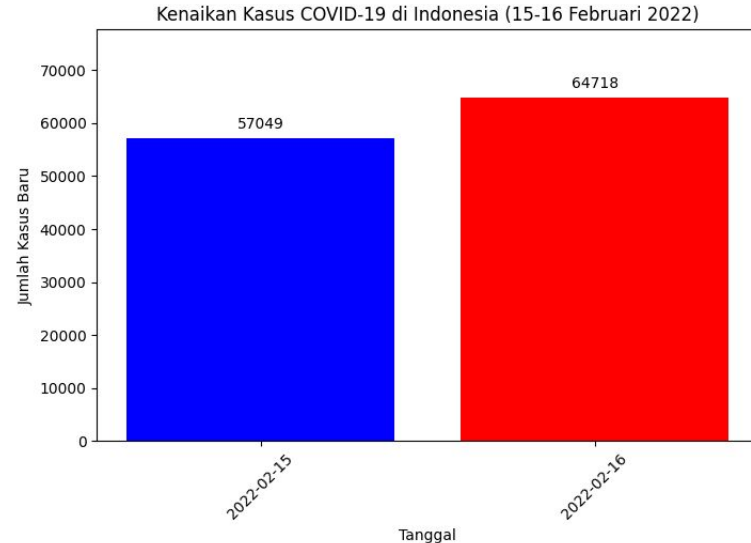
Persentase Kenaikan $= \frac{64.718 - 57.068}{57.068} \times 100\%$

$= \frac{7.650}{57.068} \times 100\% = 13.41\%$

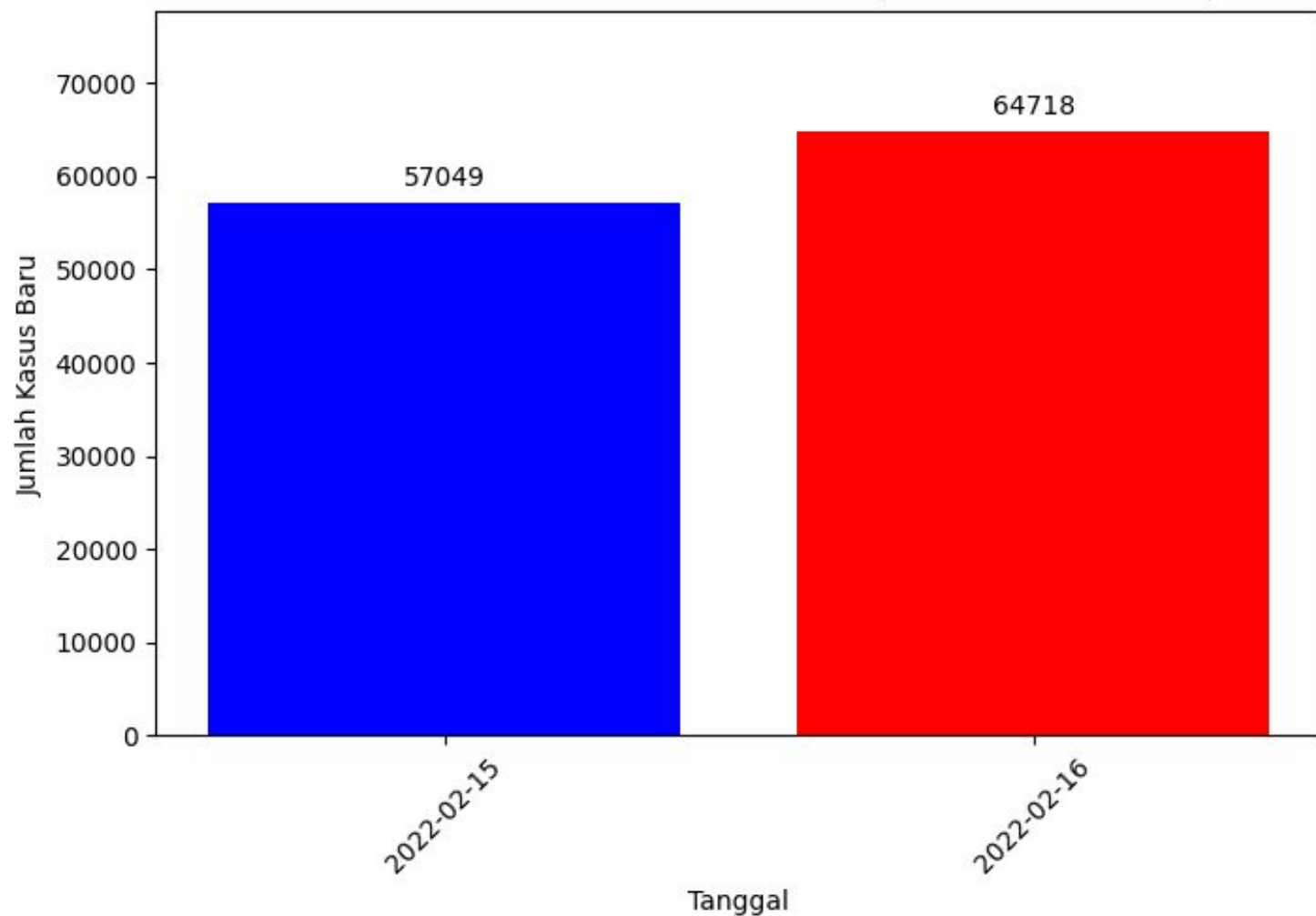
Visualisasi Data, Dari Metode Numerik - Diagram Batang

Kenaikan kasus COVID-19 di Indonesia dari 15 Februari 2022 ke 16 Februari 2022.

- Warna biru mewakili jumlah kasus pada **15 Februari (57.068 kasus)**,
- Sedangkan warna merah menunjukkan **16 Februari (64.718 kasus)**. Terjadi peningkatan **13,44%** dalam sehari, menandakan lonjakan signifikan.
- Dengan selisih kasus sebanyak, **7.650**.



Kenaikan Kasus COVID-19 di Indonesia (15-16 Februari 2022)



Kelompok 4 - Distribusi Frekuensi

File

Edit

Selection

View

Go

Run

...

ProjectPythonPengantarStatistikaKelompok

Python 3.13.2

EXPLORER

PROJ...

.ipynb_checkpoints

analysisCovid-ch...

WHO-COVID-19...

analysisCovid.ipynb

outputKasus.png

outputKasus2.png

outputKasus3.png

outputKasus4.png

outputKasus5.png

WHO-COVID-19...

analysisCovid.ipynb

stats = {

Generate

Code

Markdown

Run All

Restart

Clear All Outputs

Jupyter Variables

Outline

...

```
import numpy as np

cases_data = [prev_cases, max_case_day["New_cases"]]

interval_size = 5000
min_case = int(min(cases_data) // interval_size) * interval_size
max_case = int(max(cases_data) // interval_size + 1) * interval_size

intervals = np.arange(min_case, max_case + interval_size, interval_size)

freq, bins = np.histogram(cases_data, bins=intervals)

freq_rel = freq / sum(freq) * 100

tally_marks = ["|" * f for f in freq]

class_marks = [(bins[i] + bins[i + 1]) / 2 for i in range(len(bins) - 1)]

freq_table = pd.DataFrame({
    "Selang (Interval Kelas)": [f"{int(bins[i])} - {int(bins[i+1])}" for i in range(len(bins) - 1)],
    "Nilai (Titik Tengah)": class_marks,
    "Turus (Tally)": tally_marks,
    "Fi (Frekuensi)": freq,
    "Fri (Frekuensi Relatif, %)": freq_rel,
    "Tanda Kelas": class_marks
})

freq_table
```

[12]

Python

OUTLINE

TIMELINE

VS CODE PETS

Launchpad

0 0 0

Ln 12, Col 32

Spaces: 4

Cell 14 of 20

Subuh(04:43):-04:32:43 [KOTA JAKARTA]



Istilah-Istilah yang akan digunakan

- **Selang (Interval Kelas):** Rentang nilai kasus dalam suatu interval.
- **Nilai (Titik Tengah):** Nilai tengah dari setiap interval kelas.
- **Turus (Tally):** Representasi jumlah data dalam interval menggunakan garis turus.
- **Fi (Frekuensi):** Banyaknya data dalam suatu interval kelas.
- **Fri (Frekuensi Relatif):** Persentase data dalam interval terhadap total data.
- **Tanda Kelas:** Titik tengah dari setiap interval kelas.



Distribusi Frekuensi Kasus COVID-19 pada 15 dan 16 Februari 2022

Selang (Interval Kelas)	Nilai (Titik Tengah)	Turus (Tally)	Fi (Frekuensi)	Fri (Frekuensi Relatif, %)	Tanda Kelas
55.000 - 60.000	57.500		1	50.0	57.500
60.000 - 65.000	62.500		1	50.0	62.500



Penjelasan, Distribusi Frekuensi Kasus COVID-19 pada 15 dan 16 Februari 2022

Tanda kelas dan titik tengah dalam tabel distribusi frekuensi sering kali sama karena tanda kelas didefinisikan sebagai nilai yang mewakili suatu kelas interval.

- Titik Tengah (Nilai Kelas/Titik Kelas), didapatkan dari batas $\frac{\text{interval kelas bawah} + \text{batas interval kelas atas}}{2}$.
- Tanda Kelas, Biasanya, tanda kelas dipilih sama dengan titik tengah karena titik tengah dianggap sebagai representasi paling akurat dari seluruh nilai dalam interval tersebut.



Penjelasan, Distribusi Frekuensi Kasus COVID-19 pada 15 dan 16 Februari 2022

Ukuran interval dalam kasus ini adalah 5.000 kasus. Dengan alasan:

- **Rentang data, dipilih 5.000**, yang tidak terlalu kecil (sehingga terlalu banyak kelas) atau terlalu besar (sehingga kurang detail).
- **Standar umum, rentang data.** Dalam pembuatan histogram dan tabel distribusi frekuensi untuk membagi data menjadi kelompok-kelompok dengan lebar yang mudah diinterpretasikan.