

CYBER SECURITY DATA ANALYSIS

Laporan Tugas Besar WI2002

Mata Kuliah Literasi Data dan Intelegensi Artifisial



Disusun oleh:

Geraldo Artemius	13524005
Mikhael Adrian Yonatan	13524051
Junior Natra Situmorang	13524055
Reynard Nathanael	13524103
Nicholas Luis Chandra	13524105

SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG

April 2025

DAFTAR ISI

DAFTAR ISI.....	1
DAFTAR GAMBAR.....	4
DAFTAR TABEL.....	5
BAB 1.....	6
PENDAHULUAN.....	6
BAB 2.....	7
DESKRIPSI DATA.....	7
2.1. Atribut Data.....	7
2.1.1. Global Cybersecurity Indexes (2023).....	7
2.1.2. Cyber Security Threats (2015-2024).....	8
2.2. Regresi Data.....	9
BAB 3.....	12
PERTANYAAN PENELITIAN.....	12
BAB 4.....	13
VISUALISASI DATA.....	13
4.1. Perbandingan Kategori.....	13
4.2. Penampilan Perubahan Terhadap Waktu.....	14
4.3. Penampilan Hierarki dan Hubungan Keseluruhan-bagian.....	14
4.4. Plotting Relationships.....	16
4.5. Cara Mendapatkan Visualisasi Data.....	17
BAB 5.....	18
STATISTIK DESKRIPTIF.....	18
5.1. Dataset Global Cybersecurity Threats.....	18
5.2. Dataset Cyber Security Indexes.....	19
5.2.1. Cyber Exposure Index (CEI).....	20
5.2.2. Global Cybersecurity INdex (GCI).....	20
5.2.3. National Cyber Security Index (NCSI).....	20
5.2.4. Digital Development Level (DDL).....	21
5.3. Cara Mendapatkan Statistik Data.....	21
BAB 6.....	24
KORELASI.....	24
6.1. Pembahasan Korelasi Antar Atribut Data.....	24
6.1.1. Number of Affected Users vs Financial Loss.....	24
6.1.2. Incident Resolution Time vs Financial Loss.....	25
6.1.3. Incident Resolution Time vs Number of Affected Users.....	25
6.1.4. CEI vs NCSI.....	26
6.1.5. CEI vs GCI.....	27
6.1.6. NCSI vs GCI.....	27
6.1.7. DDL vs CEI.....	28
6.1.8. DDL vs NCSI.....	29
6.1.9. DDL vs GCI.....	29
6.2. Cara Mendapatkan Korelasi.....	30
BAB 7.....	31
DATA CLEANSING.....	31
7.1 Data Cleansing Global Cyber Security Index.....	32

7.2 Data Cleansing Dataset Global Cyber Security Threats.....	34
BAB 8.....	36
TRANSFORMASI DATA.....	36
8.1 Transformasi Dataset.....	36
8.2 Perubahan Atribut Dataset.....	36
BAB 9.....	37
DATA ANALISIS.....	37
9.1 Data Analisis Global Cybersecurity Threats.....	37
9.2 Data Analisis Global Cyber Security Indexes.....	38
KESIMPULAN.....	40

DAFTAR GAMBAR

Gambar 4.1.1. Bar Chart Count of Attack Source.....	10
Gambar 4.1.2. Bar Chart Number of Attack in A Country.....	10
Gambar 4.2.1. Line Chart Number of Attack tiap Tahun.....	11
Gambar 4.3.1. Pie Chart Count of Target Industry.....	11
Gambar 4.3.2. Pie Chart Count of Attack Type.....	12
Gambar 4.4.1. Scatter Plot Number of Affected Users vs Financial Loss.....	12
Gambar 4.4.2. CEI vs NCSI.....	13
Gambar 5.3.1. Rumus Ukuran Pemusatan (1).....	17
Gambar 5.3.2. Rumus Ukuran Pemusatan (2).....	17
Gambar 5.3.3. Rumus Ukuran Penyebaran (1).....	18
Gambar 5.3.4. Rumus Ukuran Penyebaran (2).....	18
Gambar 6.1.1. Scatter Plot Number of Affected Users vs Financial Loss.....	19
Gambar 6.1.2. Scatter Plot Incident Resolution Time vs Financial Loss.....	20
Gambar 6.1.3. Scatter Plot Incident Resolution Time vs Number of Affected Users.....	20
Gambar 6.1.4. Scatter Plot CEI vs NCSI.....	21
Gambar 6.1.5. Scatter Plot CEI vs GCI.....	21
Gambar 6.1.6. Scatter Plot NCSI vs GCI.....	22
Gambar 6.1.7. Scatter Plot DDL vs CEI.....	22
Gambar 6.1.8. Scatter Plot DDL vs NCSI.....	23
Gambar 6.1.9. Scatter Plot DDL vs GCI.....	23
Gambar 7.1.1. Code Python Data Cleansing Dataset Cyber Security Indexes.....	26
Gambar 7.1.2. Output Code Python Data Cleansing Dataset Cyber Security Indexes.....	27
Gambar 7.2.1. Code Python Data Cleansing Dataset Global Cybersecurity Threats.....	28
Gambar 7.2.2. Output Code Python Data Cleansing Dataset Global Cybersecurity Threats.....	29
Gambar 9.1.1. Code Python Data Analysis Dataset Cyber Security Indexes.....	31
Gambar 9.1.2. Output Code Python Data Analysis Dataset Cyber Security Indexes.....	32

DAFTAR TABEL

Tabel 5.1.1. Statistik Dataset Global Cybersecurity Threats.....	14
Tabel 5.2.1. Statistik Dataset Cyber Security Indexes.....	15
Tabel 11.1 Lampiran Pembagian Kerja Kelompok.....	35

BAB 1

PENDAHULUAN

Di era modern ini, informasi telah berkembang menjadi salah satu komoditas paling vital dalam berbagai bidang kehidupan, seperti kesehatan, bisnis, dan teknologi. Kemajuan ini didorong oleh peningkatan signifikan dalam digitalisasi global, sehingga informasi dan data menjadi semakin beragam serta tersedia dalam jumlah yang sangat besar. Data tidak hanya menjadi sarana untuk dokumentasi, melainkan juga dimanfaatkan secara strategis untuk menganalisis masalah dan merumuskan solusi yang efektif. Namun, seiring dengan laju digitalisasi yang pesat, tantangan baru turut bermunculan, salah satunya adalah permasalahan keamanan siber. Ancaman ini sebenarnya bukan fenomena baru, melainkan telah ada sejak era awal jaringan internet dan server berkembang. Pelaku kejahatan siber melakukan berbagai tindakan peretasan untuk mengakses data secara ilegal, melakukan pengancaman, serta menimbulkan berbagai kerugian lainnya.


Dalam melakukan implementasi literasi data, diambil dua buah dataset yang memiliki topik yang sama, yaitu *cyber security* dengan atribut data yang berbeda. Judul dari data yang diambil adalah sebagai berikut:

1. Dataset Global Cybersecurity Threats (2015-2024)

Dataset:  Dataset Global Cybersecurity Threats 2015-2024

Tipe data dan Ukuran file: xlsx (149kb)

Deskripsi Singkat: Berisi data tentang penyerangan siber yang terjadi dari tahun 2015-2024 yang dilengkapi dengan informasi pendukung

Sumber:  [Global Cybersecurity Threats \(2015-2024\)](#)

2. Dataset Cyber Security Indexes

Dataset:  Cybersecurity Index Data

Tipe data dan Ukuran file: xlsx (29.5kb)

Deskripsi Singkat: Berisi data tentang berbagai indikator yang berkaitan dengan keamanan siber relatif terhadap suatu negara pada tahun 2023.

Sumber: [Cyber Security Indexes](#)

Link Github : https://github.com/staplesmaster/TUBES_LIDIA

BAB 2

DESKRIPSI DATA

2.1. Atribut Data

Sebelum melakukan analisis lebih lanjut terhadap data-data yang terdapat di dalam dataset, pemahaman terhadap setiap atribut data yang ada sangatlah diperlukan. Berikut merupakan atribut data yang terdapat pada setiap dataset yang digunakan,

2.1.1. Global Cybersecurity Indexes (2023)

1. Country

Nama negara tempat terjadinya serangan Cyber Security yang menjadi objek pengukuran. Dataset memuat sebanyak 193 negara.

2. Region

Wilayah secara geografis negara tempat tersebut berada seperti Asia-Pasifik, Eropa, Afrika, Amerika Utara, dan Amerika Selatan.

3. CEI (Cybersecurity Exposure Index)

Indeks yang mengukur tingkat eksposur atau kerentanan sebuah negara terhadap risiko keamanan siber. Indeks berada dalam rentang 0 sampai 1, semakin tinggi indeks maka semakin tinggi *exposure*.

4. GCI (Global Cybersecurity Index)

Indeks yang menunjukkan tingkat kesiapan negara atau komitmen sebuah negara dalam menghadapi ancaman keamanan siber menurut penilaian global.

5. NCSI (National Cyber Security Index)

Indeks global yang mengukur kesiapan suatu negara dalam mencegah ancaman siber dan mengelola serangan keamanan siber.

6. DDL (Digital Development Level)

Tingkat perkembangan digital negara tersebut yang berhubungan dengan kesiapan teknologi dan penggunaan digital.

2.1.2. Cyber Security Threats (2015-2024)

1. Country

Nama negara tempat insiden keamanan siber terjadi.

2. Year

Tahun terjadinya suatu negara mengalami insiden tersebut.

3. Attack Type

Jenis serangan keamanan siber yang dilakukan, seperti *Phishing*, *Ransomware*, *Man-in-the-Middle*, *DDOS*, *SQL Injection*, dan sebagainya.

4. Target Industry

Industri atau sektor yang menjadi target dari serangan keamanan siber, misalnya Pendidikan, Retail, IT, Telekomunikasi, dan sebagainya.

5. Financial Loss

Kerugian finansial yang diakibatkan oleh serangan tersebut dan diukur dalam jutaan dolar AS.

6. Number of Affected Users

Jumlah pengguna atau individu yang menjadi korban insiden tersebut.

7. Attack Source

Sumber dari serangan keamanan siber tersebut, bisa berupa *Hacker Group*, *Nation-State*, atau *Insider*.

8. Security Vulnerability Type

Jenis kerentanan keamanan yang dimanfaatkan dalam serangan, seperti *Unpatched Software*, *Weak Passwords*, *Social Engineering*, dan lain-lain.

9. Defence Mechanism Used

Mekanisme pertahanan yang digunakan untuk menghadapi serangan keamanan siber, seperti *VPN*, *Firewall*, atau *AI-based Detection*.

10. Incident Resolution Time

Waktu yang dibutuhkan untuk menyelesaikan insiden tersebut dan diukur dalam satuan jam.

BAB 3

PERTANYAAN PENELITIAN

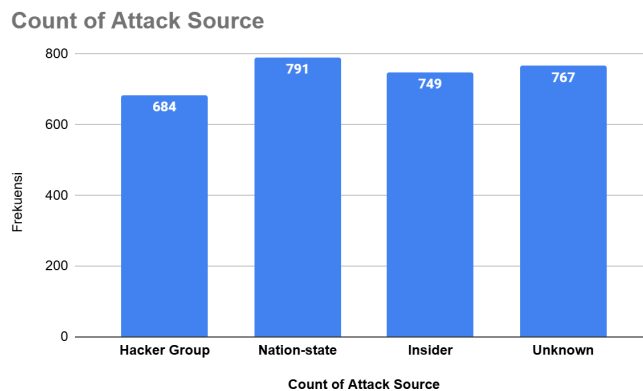
1. Bagaimana pola atau tren negara-negara di dunia dalam menghadapi dinamika ancaman siber di era digitalisasi, berdasarkan frekuensi serangan siber dan kerugian finansial yang ditimbulkan?
2. Apakah perkembangan teknologi dari tahun ke tahun berkorelasi positif dengan peningkatan indeks kesiapan keamanan siber (Cyber Security Readiness Index) serta kemampuan negara-negara dalam membangun infrastruktur digital yang lebih tahan terhadap serangan siber?
3. Bagaimana hubungan antara stabilitas politik, tingkat kemakmuran (GDP per kapita), serta tingkat inovasi teknologi (misalnya Global Innovation Index) dengan frekuensi insiden keamanan siber dan besaran kerugian finansial yang dialami negara tersebut?
4. Strategi atau langkah-langkah spesifik apa yang dapat dilakukan pemerintah suatu negara untuk meningkatkan indeks kesiapan menghadapi ancaman siber internasional, serta apakah indeks ini berkorelasi positif dengan tingkat literasi digital masyarakat negara tersebut?

BAB 4

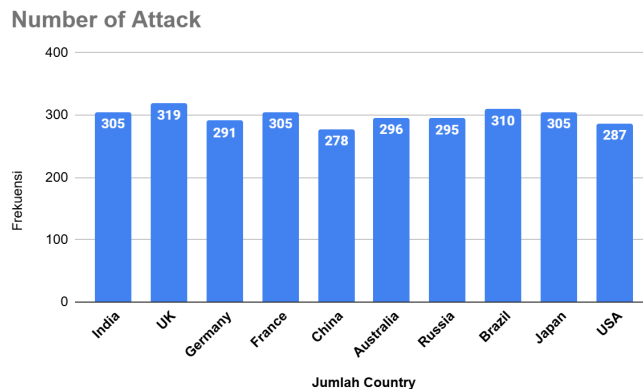
VISUALISASI DATA

4.1. Perbandingan Kategori

Bar chart digunakan untuk membandingkan kategori karena dapat menampilkan data secara jelas dan proporsional, memudahkan pembaca mengenali perbedaan antara kategori dan tren data secara efektif. *Bar chart* digunakan untuk memvisualisasikan kategori pada data pertama, yaitu Count of Attack Source dan Count of Defence Mechanism Used.



Gambar 4.1.1. *Bar Chart* Count of Attack Source



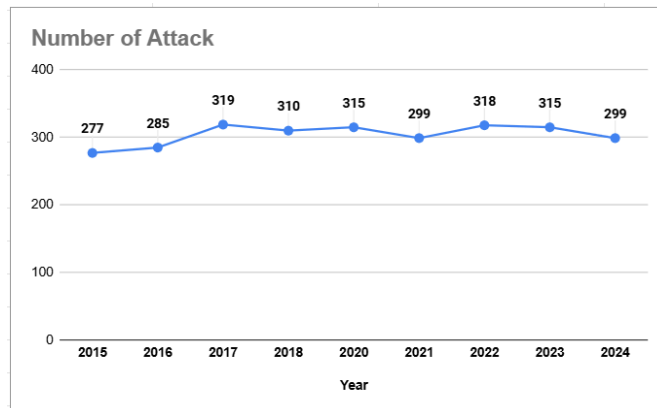
Gambar 4.1.2. *Bar Chart* Number of Attack in A Country

Dari visualisasi data gambar 4.1.1 menggunakan model visualisasi *bar chart*, terlihat bahwa sumber penyerang dengan frekuensi terbanyak berasal dari Nation-state sebanyak 791 serangan. selain itu, penyerangan terbanyak selanjutnya berasal dari sumber yang tidak dikenal (*unknown*) hal ini mengindikasi bahwa banyak penyerangan siber dilakukan oleh pihak-pihak yang tidak diketahui (*anonymous*).

Dari visualisasi data gambar 4.1.2 dapat diamati bahwa negara yang paling sering terkena penyerangan siber adalah Inggris, meskipun demikian dari 10 negara di atas dapat dilihat bahwa untuk sebagian besar negara tidak terdapat perbedaan jumlah penyerangan siber yang begitu signifikan.

4.2. Penampilan Perubahan Terhadap Waktu

Line chart digunakan untuk visualisasi data berurutan, seperti tren waktu, karena mampu menunjukkan perubahan nilai secara kontinu melalui garis yang menghubungkan titik-titik data. Dengan garis yang mengalir dari satu titik ke titik lainnya, *line chart* memudahkan identifikasi pola, trend naik atau turun, serta fluktuasi dalam rentang waktu tertentu. Visualisasi *line chart* digunakan untuk melihat Number of Attack berdasarkan urutan waktunya.

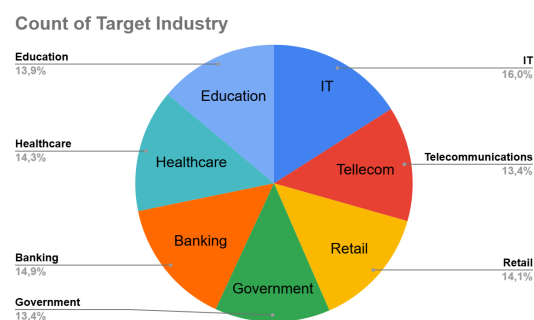


Gambar 4.2.1. *Line Chart* Number of Attack tiap Tahun

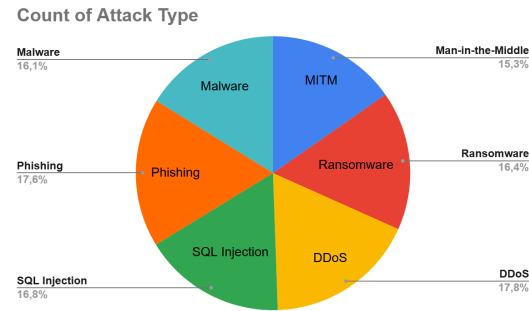
Dari visualisasi data gambar 4.2.1 dapat dilihat bahwa jumlah penyerangan siber yang terjadi dari tahun ke tahun terus mengalami kenaikan dan penurunan, meskipun demikian kejadian kenaikan jumlah penyerangan lebih sering terjadi walaupun tidak signifikan.

4.3. Penampilan Hierarki dan Hubungan Keseluruhan-bagian

Pie chart adalah salah satu jenis visualisasi data yang efektif untuk menggambarkan proporsi atau persentase dari keseluruhan. Grafik ini berfungsi dengan membagi lingkaran menjadi beberapa bagian, masing-masing mewakili kategori yang berbeda. Setiap bagian menunjukkan kontribusi relatif kategori terhadap total keseluruhan. Penggunaan *pie chart* pada dataset cyber ini, yaitu Count of Target Industry dan Count of Attack Type.



Gambar 4.3.1. *Pie Chart* Count of Target Industry



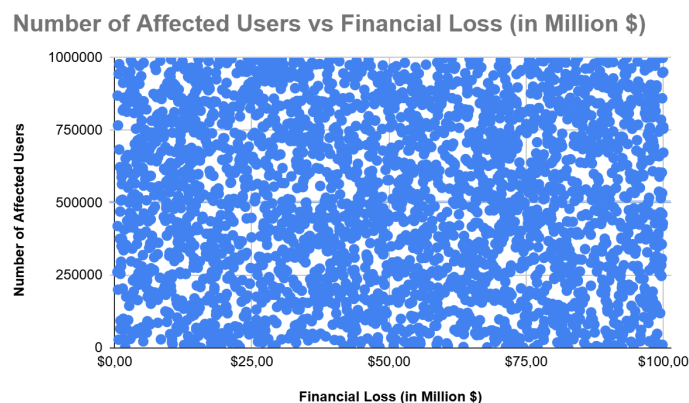
Gambar 4.3.2. *Pie Chart* Count of Attack Type

Dari visualisasi menggunakan *pie chart*, pada gambar 4.3.1. terlihat bahwa industri yang menjadi target penyerang *cyber* terbanyak berasal dari industri IT yang memiliki 16.0% dari keseluruhan target industri pada data. Hal ini menunjukkan bahwa industri IT menjadi target utama serangan siber. Industri IT memiliki risiko yang lebih tinggi terhadap serangan *cyber*.

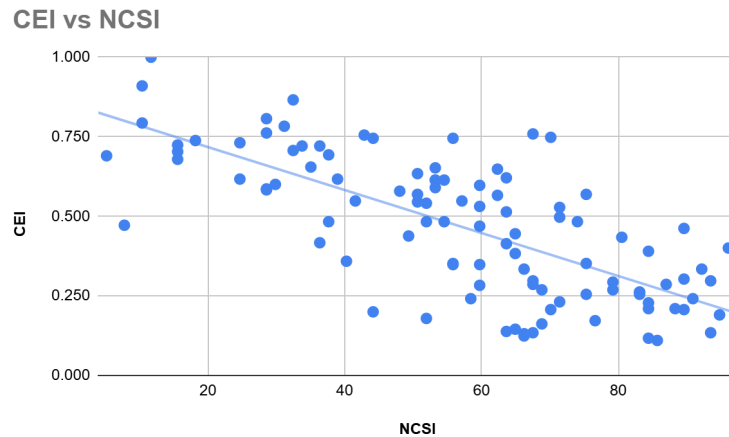
Pada gambar 4.3.2. dapat dilihat type penyerangan *cyber* terbanyak berasal dari DDoS yang memiliki 17.8% dari Seluruh tipe penyerang *cyber*. Hal ini menunjukkan bahwa serangan DDoS (Distributed Denial of Service) menjadi ancaman signifikan dalam dunia siber, mengingat kemampuannya untuk melumpuhkan sistem dengan cara membanjiri jaringan, server, atau layanan dengan lalu lintas data yang berlebihan, sehingga mengganggu akses pengguna yang sah.

4.4. Plotting Relationships

Scatter plot adalah jenis visualisasi data yang digunakan untuk menampilkan hubungan antara dua variabel kuantitatif melalui titik-titik data pada bidang kartesius. Setiap titik merepresentasikan pasangan nilai dari dua variabel tersebut, memungkinkan pengamatan pola distribusi, korelasi, atau tren di antara keduanya. penggunaan *scatter plot* pada dataset *cyber* ini, yaitu Number of Affected Users vs Financial Loss dan CEI vs NCSI.



Gambar 4.4.1. *Scatter Plot* Number of Affected Users vs Financial Loss



Gambar 4.4.2. CEI vs NCSI

Dari visualisasi data gambar 4.4.1 yang membandingkan dua variabel, yaitu jumlah pengguna terdampak dengan kerugian dapat dilihat bahwa scatter plot menampilkan pola yang sangat acak. Maka, dapat disimpulkan tidak terdapat korelasi antara jumlah pengguna terdampak dengan kerugian.

Pada visualisasi data gambar 4.4.2 membandingkan dua variabel juga, yaitu CEI dengan NCSI. CEI adalah indeks yang menjadi indikator seberapa sering suatu negara terkena serangan siber, sementara NCSI adalah indeks yang mengukur seberapa siap suatu negara dalam menghadapi serangan siber. Visualisasi data di atas menunjukkan korelasi negatif antara CEI dengan NCSI yang memiliki arti bahwa negara yang dinilai siap menghadapi serangan siber (NCSI tinggi) cenderung mendapatkan serangan siber yang lebih sedikit (CEI rendah).

4.5. Cara Mendapatkan Visualisasi Data

Untuk mendapatkan visualisasi data di atas, digunakan *tool* yang telah disediakan oleh Google Spreadsheet. Adapun cara yang dilakukan adalah melakukan *select* terhadap seluruh data dalam suatu *column* (kategori yang dipilih), kemudian membuka menu *insert* dan memilih menu diagram. Setelah diagram berhasil dibuat, dapat dilakukan pengaturan terhadap jenis diagram yang diinginkan (*bar chart*, *line chart*, *pie chart*, *scatter plot*, dsb), nama judul, nama label pada sumbu xy, menampilkan label frekuensi pada diagram, ukuran *font*, jenis *font*, dan pewarnaan.

Namun, dalam Google Spreadsheet terdapat beberapa “catatan khusus” untuk jenis diagram tertentu. Untuk jenis diagram *scatter plot* perlu dilakukan *select* terhadap seluruh data dari dua kolom (dua kategori) yang ingin dibandingkan. Kemudian untuk data yang butuh diurutkan seperti data waktu (untuk melihat tren dari waktu ke waktu), pengurutan tidak dapat dilakukan pada pengaturan diagram karena tidak ada fitur demikian sehingga harus dilakukan pengurutan secara langsung pada tabel data.

BAB 5

STATISTIK DESKRIPTIF

5.1. Dataset Global Cybersecurity Threats

Berdasarkan dataset Global Cybersecurity Threats, didapatkan statistik sebagai berikut:

	Financial Loss (in million \$)	Number of Affected Users	Incident Resolution Time (in Hours)
Rata-Rata	50,92	504150,345	36
Maksimum	99,99	999.635,00	72
Minimum	0,50	6,00	1
Persentil 10%	10,82	95.351,00	8
Kuartil 1	25,78	7.801.378.960,50	968
Median	50,90	503775	37
Kuartil 3	75,63	290.297,91	21
Persentil 90%	89,79	902.112,20	65
Variansi	828,83	84.244.701.926,23	423
Standar Deviasi	28,79	290.249,38	21

Tabel 5.1.1. Statistik Dataset Global Cybersecurity Threats

Dari perhitungan statistik pada tabel di atas, dapat diperoleh suatu informasi pada setiap kategori yang bersangkutan. Pada kerugian finansial, diperoleh perhitungan standar deviasi sebesar 28,79 juta dolar AS yang mengindikasikan bahwa penyebaran data cukup besar karena didapatkan koefisien variansi (CV) sebesar 55% dimana CV lebih besar dari 30% menandakan bahwa variabilitas data tinggi. Hal itu menunjukkan bahwa besarnya kerugian finansial akibat serangan siber sangat bervariasi dan tidak konsisten antar kejadian. Selain itu, dapat diketahui bahwa rata-rata kerugian finansial dari semua penyerangan siber sebesar 50, 92 juta dolar AS dengan nilai median pada 50, 9 juta dolar AS. Karena nilai mean dan median hampir sama, dapat disimpulkan bahwa distribusi data cukup simetris. Hal tersebut mengindikasikan bahwa tidak ada pencilan besar yang mendistorsi nilai rata-rata secara signifikan.

Begitu pula pada jumlah pengguna yang terdampak dan waktu resolusi tiap insiden, keduanya memiliki sifat yang sama dengan data kerugian finansial yaitu memiliki standar deviasi yang tergolong

besar dan nilai mean dengan median yang relatif sama (simetris). Hal tersebut menandakan bahwa penyebaran data cukup besar, namun tidak ada pencilon besar yang mendistorsi nilai rata-rata secara signifikan.

5.2. Dataset Cyber Security Indexes

Berdasarkan dataset Cyber Security Indexes, didapatkan statistik sebagai berikut:

	CEI	GCI	NCSI	DDL
Rata-Rata	0.463	74.23	58.13	57.05
Maksimum	1.000	100.000	96.100	82.930
Minimum	0.110	2.200	10.390	19.500
Persentil 1	0.181	27.948	28.570	33.346
Kuartil 1	0.269	64.020	42.210	45.390
Median	0.483	84.760	59.740	58.870
Kuartil 3	0.617	96.190	74.675	69.475
Persentil 9	0.744	97.908	86.750	78.680
Variansi	0.046	727.719	487.971	274.585
Standar Deviasi	0.215	26.976	22.090	16.571

Tabel 5.2.1. Statistik Dataset Cyber Security Indexes

5.2.1. Cyber Exposure Index (CEI)

Nilai CEI memiliki rata-rata 0,463, menunjukkan bahwa rata-rata paparan siber negara-negara berada pada tingkat sedang. Dengan nilai maksimum 1,000 menunjukkan setidaknya ada negara yang memiliki paparan siber sangat tinggi. Minimum sebesar 0,110 mengindikasikan adanya negara dengan paparan siber sangat rendah. Median sebesar 0,483 menunjukkan distribusi data relatif simetris karena dekat dengan rata-rata (0,463). Standar deviasi sebesar 0,215 menunjukkan tingkat variasi data moderat, artinya negara-negara memiliki perbedaan paparan yang cukup nyata namun tidak terlalu ekstrem.

5.2.2. Global Cybersecurity Index (GCI)

Rata-rata GCI sebesar 74,23 dari maksimum 100, menandakan rata-rata negara memiliki tingkat keamanan siber yang cukup tinggi. Minimum yang rendah (2,200)

menunjukkan ada negara dengan kondisi keamanan siber yang sangat buruk. Median (84,760) lebih besar dari rata-rata (74,23) menunjukkan sebagian besar negara memiliki nilai yang tinggi (distribusi sedikit skewed ke kiri atau negatif), artinya banyak negara relatif memiliki keamanan siber yang baik, dengan sedikit negara memiliki nilai rendah yang menyebabkan rata-rata menurun. Variansi yang besar (727,719) dan standar deviasi tinggi (26,976) menunjukkan adanya ketimpangan yang signifikan antara negara dengan keamanan tinggi dan rendah.

5.2.3. National Cyber Security Index (NCSI)

Nilai rata-rata NCSI (58,13) menunjukkan tingkat kesiapan nasional menghadapi ancaman siber yang moderat. Nilai maksimum sebesar 96,100 menunjukkan beberapa negara sangat siap menghadapi ancaman siber internasional, sementara nilai minimum (10,390) menandakan negara yang sangat tertinggal dalam kesiapan. Median sebesar 59,740 sedikit lebih tinggi dari rata-rata (58,13), menunjukkan distribusi data yang cenderung simetris namun ada sedikit skew ke kiri (negatif), menandakan sebagian besar negara memiliki kesiapan di atas rata-rata. Standar deviasi yang tinggi (22,090) menandakan variasi kesiapan antar-negara cukup besar.

5.2.4. Digital Development Level (DDL)

Nilai rata-rata DDL sebesar 57,05 mengindikasikan perkembangan digital yang relatif moderat. Rentang antara nilai maksimum (82,930) dan minimum (19,500) cukup lebar, mencerminkan adanya kesenjangan signifikan dalam perkembangan digital antar negara. Median sebesar 58,870 mendekati rata-rata, menunjukkan distribusi data cenderung simetris. Standar deviasi (16,571) yang cukup besar menandakan variasi tingkat perkembangan digital cukup nyata antar negara.

5.3. Cara Mendapatkan Statistik Data

Sebelum mendapatkan statistik dari kedua dataset, perlu dilakukan data preparation terlebih dahulu. Jika terdapat atribut dari data yang hilang atau *missing data* maka data tersebut akan dihapus untuk memudahkan perhitungan statistik dataset. Selain itu, pada tahap data preparation juga dilakukan normalisasi untuk menyelaraskan format dari dataset. Misalnya dengan membuat seluruh data pada kolom 'A' menjadi bentuk data *number* dan mengubah tanda koma menjadi tanda titik.



Ukuran Pemusatan

Statistik yang memberikan informasi terkait tempat data terkumpul dengan ukuran/jumlah tertentu

n datum pengamatan/observasi:

$$o_1, o_2, \dots, o_n$$

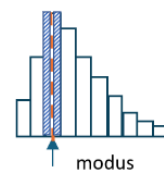
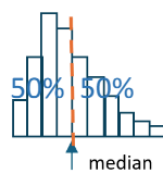
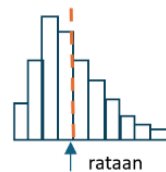


Data terurut:

$$o_{(1)}, o_{(2)}, \dots, o_{(n)}$$

*memakai kurung

- Mean (rerata) : $\bar{o} = \frac{1}{n} \sum_{i=1}^n o_i$ (data tidak perlu diurutkan)
- Median (kuartil tengah) : $q_2 = o_{\left(\frac{n+1}{2}\right)}$ (data perlu diurutkan)
- Modus, yaitu nilai yang paling sering muncul



Gambar 5.3.1. Rumus Ukuran Pemusatan (1)

- Kuartil: lokasi datum dengan pembagian data menjadi 4 bagian
 - Kuartil bawah : $q_1 = o_{\left(\frac{n+1}{4}\right)}$
 - Kuartil tengah (median) : $q_2 = o_{\left(\frac{n+1}{2}\right)}$
 - Kuartil atas : $q_3 = o_{\left(\frac{3(n+1)}{4}\right)}$
- Desil ke-k : lokasi datum dengan pembagian data menjadi 10 bagian

$$d_k = o_{\left(\frac{k(n+1)}{10}\right)}, k = 1, 2, \dots, 9$$
- Persentil ke-k : lokasi datum dengan pembagian data menjadi 100 bagian
 - Umumnya untuk data yang banyak (lebih dari 100)

$$p_k = o_{\left(\frac{k(n+1)}{100}\right)}, k = 1, 2, \dots, 99$$
- Minimum dan maksimum : nilai datum terendah dan tertinggi

$$o_{(1)} = \min \text{ dan } o_{(n)} = \max$$

Gambar 5.3.2. Rumus Ukuran Pemusatan (2)



Ukuran Penyebaran

Statistik yang menyatakan sejauh mana data dalam suatu kumpulan nilai **tersebar** atau bervariasi di sekitar rerata (mean) atau titik pusat lainnya. Penyebaran menggambarkan tingkat variasi atau keragaman dalam data.

- Variansi :

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (o_i - \bar{o})^2$$

Dalam perhitungan, s^2 dapat dihitung:

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^n o_i^2 - \frac{(\sum_{i=1}^n o_i)^2}{n} \right] \text{ atau } s^2 = \frac{1}{n-1} [\sum_{i=1}^n o_i^2 - n\bar{o}^2]$$

**rumus di atas lebih mudah dan mengurangi banyaknya pembulatan, ini yang dipakai ketika memiliki sampel.*

- Variansi yang *bias* :

$$s_n^2 = \frac{1}{n} \left[\sum_{i=1}^n o_i^2 - \frac{(\sum_{i=1}^n o_i)^2}{n} \right]$$



Gambar 5.3.3. Rumus Ukuran Penyebaran (1)



Ukuran Penyebaran (2)

- Simpangan baku (standar deviasi) :

$$s = \sqrt{s^2}$$

- Range (Jangkauan) :

$$J = maks - min = o_{(n)} - o_{(1)}$$

- Jangkauan inter-kuartil :

$$dq = q_3 - q_1$$

- Koefisien Variasi :

$$CV = \frac{s}{\bar{o}} \times 100\%$$

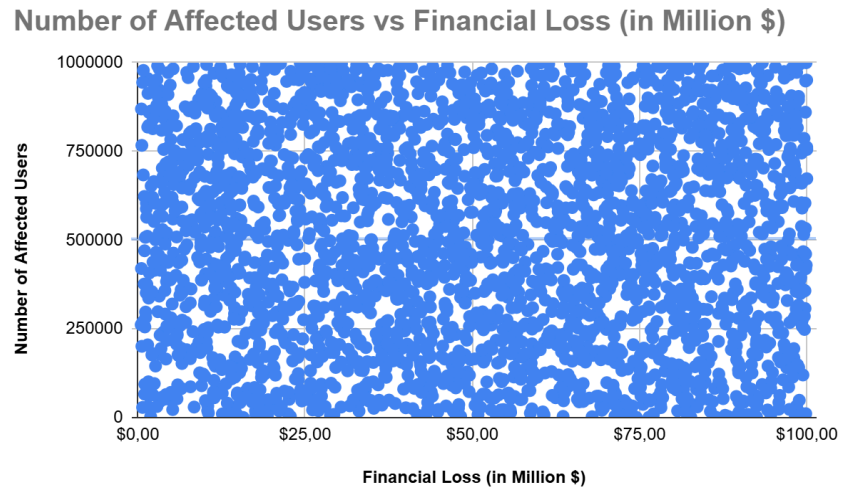
Gambar 5.3.4. Rumus Ukuran Penyebaran (2)

BAB 6

KORELASI

6.1. Pembahasan Korelasi Antar Atribut Data

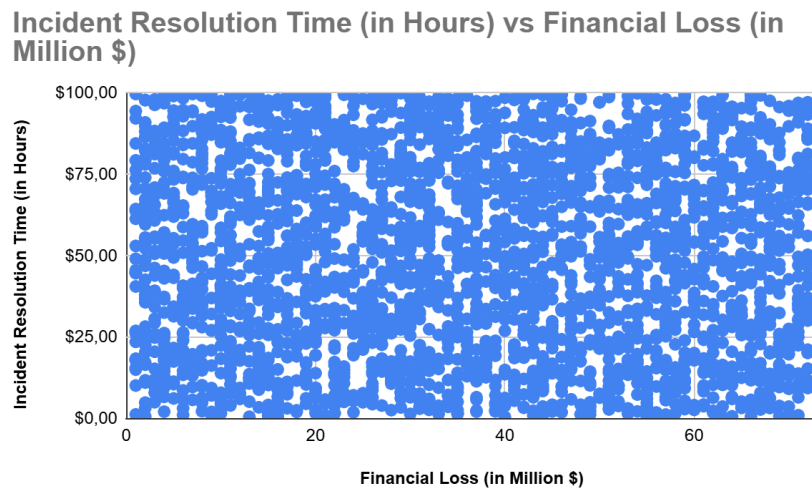
6.1.1. Number of Affected Users vs Financial Loss



Gambar 6.1.1. *Scatter Plot* Number of Affected Users vs Financial Loss

Berdasarkan *scatter plot* di atas, tidak terdapat korelasi antara *number of affected user* dengan *financial loss*. Hal tersebut menunjukkan bahwa jumlah pengguna yang terdampak tidak selalu mempengaruhi jumlah kerugian dari suatu kejadian penyerangan siber, begitu juga sebaliknya. Misalnya, terdapat suatu serangan siber yang berdampak bagi banyak pengguna, namun kerugiannya kecil karena penyerangan siber dilakukan pada bidang yang “tidak relevan” (contohnya bidang pendidikan yang tidak melibatkan keuangan). Akan tetapi, terdapat juga suatu serangan siber yang berdampak bagi banyak pengguna, namun kerugiannya besar karena penyerangan siber dilakukan pada bidang yang “relevan” (contohnya perbankan yang sangat erat dengan uang).

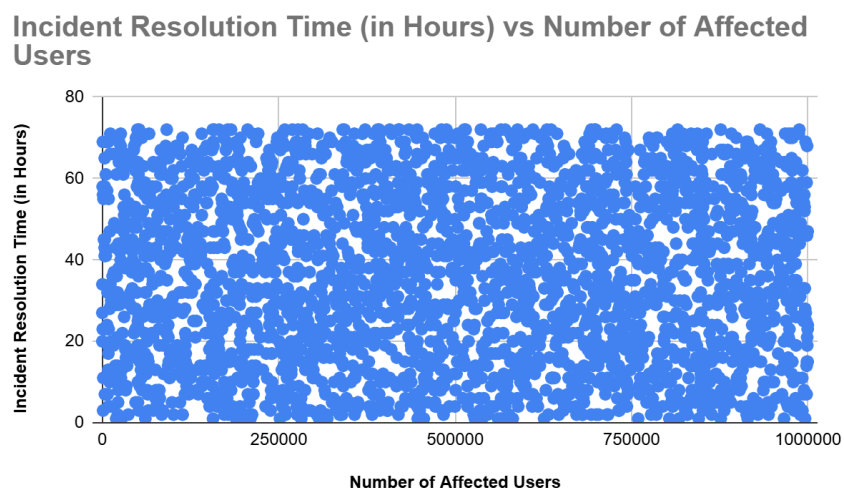
6.1.2. Incident Resolution Time vs Financial Loss



Gambar 6.1.2. *Scatter Plot* Incident Resolution Time vs Financial Loss

Berdasarkan *scatter plot* di atas, tidak terdapat korelasi antara *incident resolution time* dengan *financial loss*. Hal tersebut menunjukkan bahwa durasi penyelesaian insiden tidak selalu mempengaruhi atau dipengaruhi jumlah kerugian dari suatu kejadian penyerangan siber, begitu juga sebaliknya sehingga dapat diketahui bahwa salah satu hal yang menentukan besar dan kecilnya *financial loss* dan *incident resolution time* adalah jenis serangan siber itu sendiri yang dapat ditinjau dari kekompleksan serangan, bidang yang diserang, dan sebagainya.

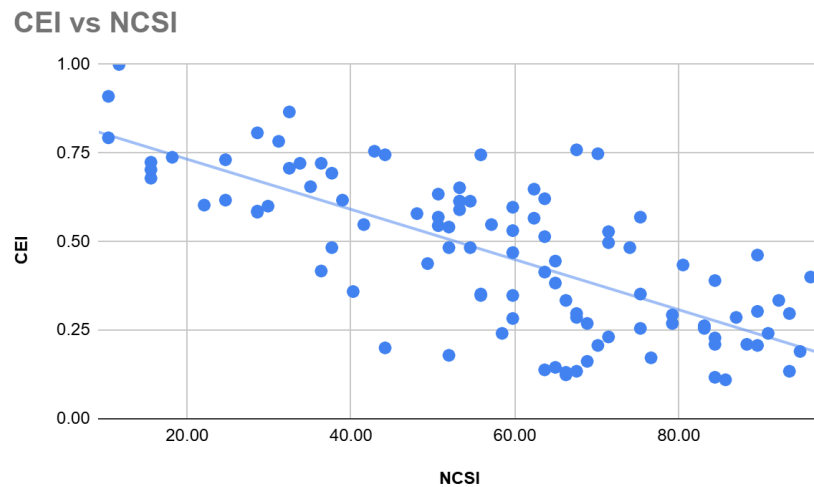
6.1.3. Incident Resolution Time vs Number of Affected Users



Gambar 6.1.3. *Scatter Plot* Incident Resolution Time vs Number of Affected Users

Tidak ada korelasi antara *Incident Resolution Time* dengan *Number of Affected Users*. Hal tersebut membuktikan bahwa waktu untuk mengatasi sebuah keamanan siber tergantung pada kompleksitas serangan, jenis ancaman, dan keahlian seorang *Cybersecurity*.

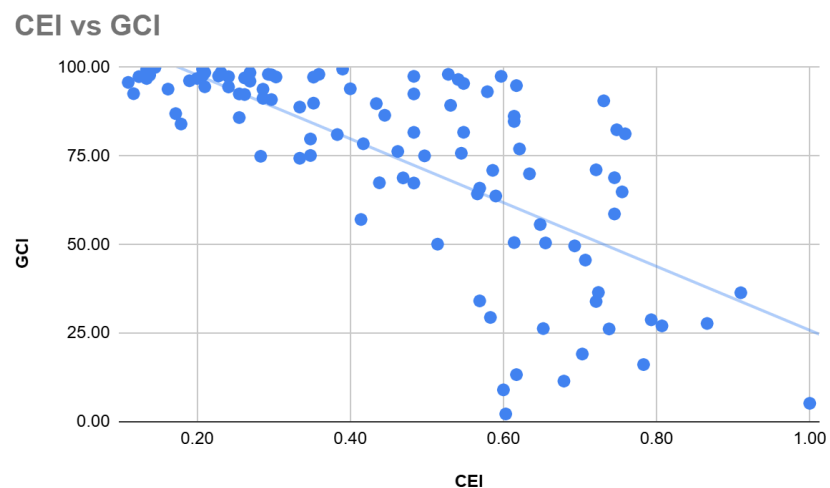
6.1.4. CEI vs NCSI



Gambar 6.1.4. *Scatter Plot* CEI vs NCSI

Korelasi antara indeks CEI dan NCSI bernilai -0.729 (korelasi negatif). CEI mengukur kerentanan ancaman siber, sementara NCSI mengukur tingkat kesiapan sebuah negara dalam menghadapi ancaman siber. Korelasi di atas menunjukkan bahwa semakin baik kesiapan suatu negara (NCSI), maka semakin kecil kemungkinan sistemnya terekspos serangan (CEI) dan berlaku sebaliknya.

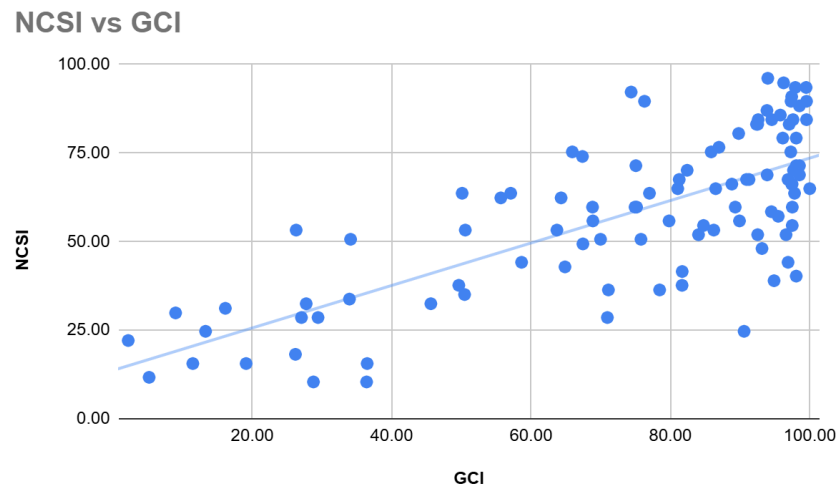
6.1.5. CEI vs GCI



Gambar 6.1.5. *Scatter Plot* CEI vs GCI

Korelasi antara indeks CEI dan GCI bernilai -0.72 (korelasi negatif). CEI mengukur kerentanan ancaman siber, sementara GCI mengukur komitmen sebuah negara dalam menghadapi ancaman siber. Korelasi di atas menunjukkan bahwa semakin tinggi komitmen suatu negara dalam menghadapi serangan siber (GCI), maka semakin tidak rentan terhadap serangan siber (CEI) dan berlaku sebaliknya.

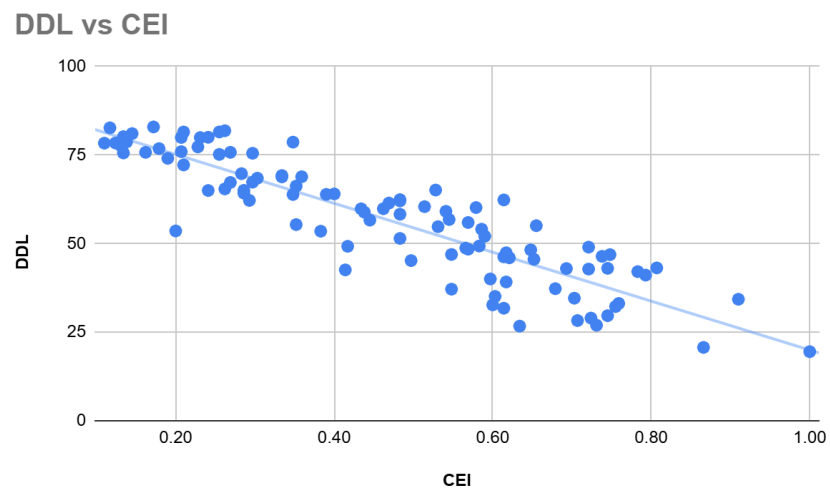
6.1.6. NCSI vs GCI



Gambar 6.1.6. *Scatter Plot* NCSI vs GCI

Korelasi antara indeks NCSI dan GCI bernilai 0.7318519471 (korelasi positif). NCSI mengukur tingkat kesiapan sebuah negara dalam menghadapi ancaman siber, sementara GCI mengukur komitmen sebuah negara dalam menghadapi ancaman siber. Korelasi di atas menunjukkan bahwa semakin tinggi komitmen suatu negara dalam menghadapi serangan siber (GCI), maka semakin tinggi kesiapan suatu negara dalam menghadapi serangan siber (NCSI) dan berlaku sebaliknya.

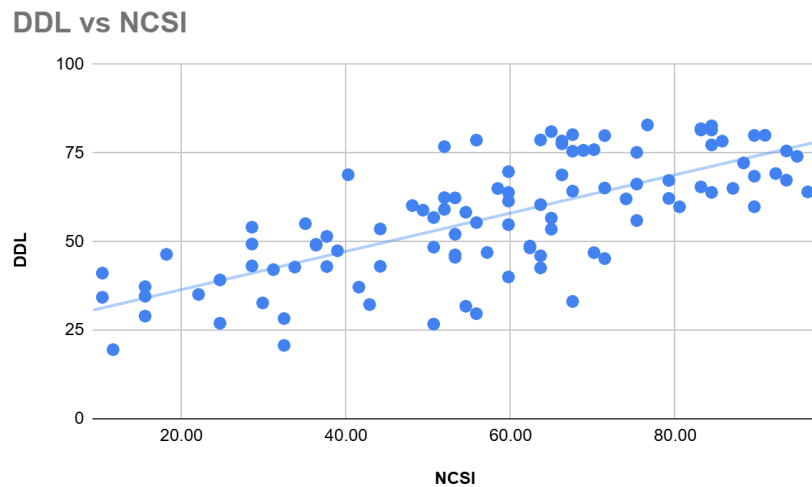
6.1.7. DDL vs CEI



Gambar 6.1.7. *Scatter Plot* DDL vs CEI

Didapatkan korelasi antara DDL dengan CEI sebesar -0.8941431885 (korelasi negatif) yang menunjukkan bahwa negara dengan DDL rendah biasanya belum memiliki kesiapan atau kesadaran keamanan digital yang memadai sehingga akan lebih rentan terhadap serangan keamanan siber (CEI).

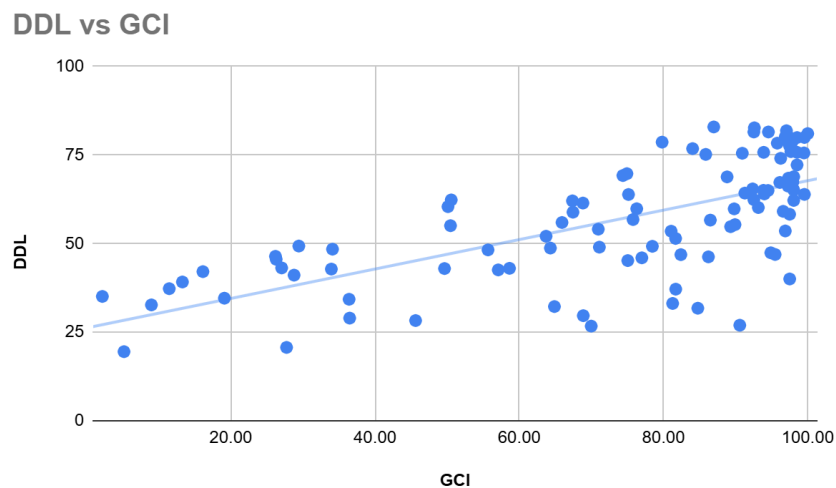
6.1.8. DDL vs NCSI



Gambar 6.1.8. *Scatter Plot* DDL vs NCSI

Didapatkan korelasi antara DDL dengan NCSI sebesar yang 0.71996734 (korelasi positif) menunjukkan hubungan positif yang cukup kuat, artinya negara dengan ketergantungan digital yang lebih tinggi cenderung memiliki tingkat kesiapan keamanan siber yang lebih baik. Semakin tinggi ketergantungan suatu negara pada teknologi (DDL), maka semakin tinggi upaya untuk memperkuat pertahanan siber (NCSI).

6.1.9. DDL vs GCI



Gambar 6.1.9. *Scatter Plot* DDL vs GCI

Didapatkan korelasi antara DDL dan GCI sebesar 0.68 menunjukkan hubungan korelasi positif karena semakin tinggi tingkat kesiapan teknologi dan penggunaan digital sebuah negara maka biasanya memiliki tingkat kesiapan keamanan siber yang tinggi juga. Hal ini juga membuktikan bahwa kesadaran akan ancaman siber meningkat seiring dengan pertumbuhan ketergantungan pada teknologi.

6.2. Cara Mendapatkan Korelasi

Pada setiap korelasi antar atribut data terdapat diagram yang berupa *scatter plot* dan suatu angka yang disebut koefisien korelasi. Cara yang kami gunakan untuk mendapatkan kedua hal tersebut adalah sebagai berikut,

Untuk mendapatkan *scatter plot*, digunakan *tool* yang telah disediakan oleh Google Spreadsheet. Adapun cara yang dilakukan adalah melakukan *select* terhadap seluruh data pada dua *column* atribut data yang ingin dibandingkan, kemudian membuka menu *insert* dan memilih menu *diagram*. Setelah diagram berhasil dibuat, dapat dilakukan pengaturan terhadap jenis diagram yang diinginkan (pada bagian ini dipilih jenis diagram *scatter plot*), nama judul, nama label pada sumbu xy, ukuran *font*, jenis *font*, dan pewarnaan.

Kemudian, untuk menentukan koefisien korelasi digunakan juga *tool* yang telah disediakan oleh Google Spreadsheet. Google Spreadsheet menyediakan banyak *tool* untuk membersihkan dan menghitung data dengan formula-formula tertentu. Hal tersebut memudahkan kami dalam mendapatkan data secara kualitatif maupun kuantitatif. Formula yang digunakan untuk menentukan koefisien korelasi adalah *correl(data y, data x)*.

BAB 7

DATA CLEANSING

Data cleansing merupakan tahap penting dalam proses pemrosesan data yang bertujuan untuk memastikan bahwa data yang digunakan bersih, konsisten, dan layak untuk dianalisis lebih lanjut. Dalam tahap ini, dilakukan identifikasi dan penanganan terhadap data yang tidak valid atau berpotensi mengganggu hasil analisis. Salah satu langkah awal yang dilakukan adalah menghapus baris yang memiliki atribut bernilai kosong (missing values). Keberadaan nilai kosong dapat menurunkan kualitas hasil analisis dan menyebabkan error pada algoritma pemrosesan data, sehingga harus dihapus untuk menjaga integritas dataset.

Langkah berikutnya dalam proses pembersihan adalah menghapus data yang duplikat. Data duplikat merupakan baris-baris dalam dataset yang memiliki isi identik dengan baris lainnya. Duplikasi semacam ini bisa muncul karena kesalahan dalam proses input data atau penggabungan data dari berbagai sumber. Jika tidak ditangani, data duplikat dapat menyebabkan bias dalam analisis karena memberikan bobot berlebih terhadap informasi tertentu. Pembersihan pada tahun-tahun yang tidak logis (lebih dari tahun 2025) dan penghapusan baris dan kolom yang hilang.

7.1 Data Cleansing Global Cyber Security Index

Code Python Data Cleansing untuk Dataset_Cyber_Security_Indexes_Noisy.csv adalah sebagai berikut.

```
import pandas as pd

# Load dataset
file_path = "../Dataset/Dataset_Cyber_Security_Indexes_Noisy.csv"
df = pd.read_csv(file_path)

# Simpan jumlah baris awal
jumlah_awal = len(df)

# Hapus duplikat
df = df.drop_duplicates()
jumlah_setelah_dedup = len(df)
jumlah_duplikat_dihapus = jumlah_awal - jumlah_setelah_dedup

# Hapus baris yang mengandung nilai kosong (NA)
df = df.dropna()
jumlah_akhir = len(df)
jumlah_na_dihapus = jumlah_setelah_dedup - jumlah_akhir

# Hitung total baris yang dihapus
total_dihapus = jumlah_duplikat_dihapus + jumlah_na_dihapus

# Tampilkan ringkasan hasil
print("=" * 40)
print("Ringkasan Pembersihan Dataset")
print("=" * 40)
print(f"Total baris awal                : {jumlah_awal}")
print(f"Duplikat yang dihapus            : {jumlah_duplikat_dihapus}")
print(f"Baris dengan NA yang dihapus     : {jumlah_na_dihapus}")
print(f"Total baris akhir                : {jumlah_akhir}")
print(f"Total baris yang dihapus         : {total_dihapus}")
print("=" * 40)

# Simpan dataset yang telah dibersihkan
output_path = "../Dataset/Dataset_Cyber_Security_Indexes_Cleaned.csv"
df.to_csv(output_path, index=False)
```

Gambar 7.1.1. Code Python Data Cleansing Dataset Cyber Security Indexes

Output yang dihasilkan:

```
=====
Ringkasan Pembersihan Dataset
=====
Total baris awal           : 192
Duplikat yang dihapus      : 0
Baris dengan NA yang dihapus : 88
Total baris akhir          : 104
Total baris yang dihapus    : 88
=====
```

Gambar 7.1.2. Output Code Python Data Cleansing Dataset Cyber Security Indexes

Dari hasil proses *data cleansing* yang telah dilakukan, diketahui bahwa terdapat 88 baris data yang memiliki nilai kosong (*missing/NA*). Baris-baris tersebut dihapus untuk memastikan integritas dan kelengkapan data yang akan dianalisis. Setelah penghapusan baris dengan nilai kosong tersebut, jumlah data yang tersisa adalah 104 baris dari total awal 192 baris. Langkah ini penting agar analisis selanjutnya tidak terpengaruh oleh data yang tidak lengkap, yang dapat menyebabkan bias atau hasil yang tidak akurat.

7.2 Data Cleansing Dataset Global Cyber Security Threats

Code Python Data Cleansing untuk Dataset_Global_Cybersecurity_Threats.csv, sebagai berikut.

```
import pandas as pd
import os

# Load data
df = pd.read_csv("../Dataset/Dataset_Global_Cybersecurity_Threats_Noisy.csv")

# Simpan jumlah awal
initial_rows = len(df)

# Hapus duplikat
df = df.drop_duplicates()
after_dedup_rows = len(df)
deleted_duplicates = initial_rows - after_dedup_rows

# Hapus baris yang mengandung NA
df = df.dropna()
after_deNA_rows = len(df)
deleted_na = after_dedup_rows - after_deNA_rows

# Pastikan kolom Year numerik
df['Year'] = pd.to_numeric(df['Year'], errors='coerce')

# Hapus baris dengan Year > 2025
df = df[df['Year'] <= 2025]
final_rows = len(df)
deleted_Years = after_deNA_rows - final_rows

# Total yang dihapus
total_deleted = deleted_duplicates + deleted_na + deleted_Years

# Cetak hasil
print("=" * 40)
print("Ringkasan Pembersihan Dataset")
print("=" * 40)
print(f"Total awal: {initial_rows}")
print(f"Duplikat yang dihapus: {deleted_duplicates}")
print(f"Baris yang mengandung NA dihapus: {deleted_na}")
print(f"Baris dengan Year > 2025: {deleted_Years}")
print(f"Total baris akhir: {final_rows}")
print(f"Total baris yang dihapus: {total_deleted}")
print("=" * 40)

# Simpan hasil
os.makedirs('Dataset', exist_ok=True)
df.to_csv('../Dataset/Dataset_Global_Cybersecurity_Threats_Cleaned.csv', index=False)
df.to_csv('../Dataset/Dataset_Global_Cybersecurity_Threats_Cleaned.csv', index=False)
```

Gambar 7.2.1. Code Python Data Cleansing Dataset Global Cybersecurity Threats

Output yang dihasilkan:

```
=====
Ringkasan Pembersihan Dataset
=====
Total awal: 3000
Duplikat yang dihapus: 0
Baris yang mengandung NA dihapus: 1890
Baris dengan Year > 2025: 102
Total baris akhir: 1008
Total baris yang dihapus: 1992
=====
```

Gambar 7.2.2. Output Code Python Data Cleansing Dataset Global Cybersecurity Threats

Berdasarkan hasil dari proses *data cleansing*, diperoleh informasi bahwa jumlah data awal adalah 3000 baris. Setelah dilakukan penghapusan terhadap data yang tidak layak, didapatkan hasil sebagai berikut. Secara keseluruhan, total baris yang dihapus mencapai 1992 baris dari 3000 data awal. Laporan pembersihan data mengungkapkan bahwa dari 3000 data awal, sebanyak 1992 baris dihapus karena nilai kosong (1890 baris) dan tahun tidak valid (102 baris melebihi 2025), tanpa adanya duplikasi, sehingga tersisa 1008 baris data yang siap untuk dianalisis.

BAB 8

TRANSFORMASI DATA

8.1 Transformasi Dataset

Pada dataset yang dianalisis didapatkan bahwa terdapat perbedaan satuan dan rentang nilai. Dataset Global Cybersecurity Threats (2015-2024) memiliki perbedaan satuan yang berupa *currency* dengan rentang nilai mulai dari 0 hingga tak terhingga, satuan number dengan rentang nilai mulai dari 0 hingga 1.000.000, dan satuan plainteks. Atribut data pada Dataset Global Cybersecurity Threats (2015-2024) memiliki nilai dan rentang nilai yang berbeda-beda. Pada Dataset Cyber Security Indexes terdapat perbedaan satuan plainteks dengan satuan *number* pada rentang nilai 0-1 dan 0-100.

8.2 Perubahan Atribut Dataset

Meskipun kedua dataset memiliki perbedaan, dataset yang digunakan tidak diberlakukan perubahan atau transformasi apapun. Hal ini dilakukan karena data atau atribut yang terkandung di dalam kedua dataset bersifat absolut atau memiliki makna tersendiri yang tidak bisa ditransformasikan ke bentuk atau satuan lain. Misalnya, nama negara tidak mungkin digantikan menjadi angka dalam dataset, sehingga data serta atribut yang sudah terkandung di dalam kedua dataset bersifat tetap dan tidak berlakukan transformasi.

BAB 9

DATA ANALISIS

Analisis data sederhana menggunakan program sederhana dalam bahasa python untuk menghasilkan model regresi linear. Masing-masing dataset memiliki variabel dependen dan variabel independennya masing-masing.

9.1 Data Analisis Global Cybersecurity Threats

Dalam dataset Global Cyber Security Threats, variabel yang digunakan meliputi *number of affected user* sebagai variabel dependen dan *financial loss* sebagai variabel independen. Berikut merupakan kode yang digunakan untuk mencari regresi linier, beserta dengan keluaran dari program.

```
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats

# Load CSV dengan penanganan koma dan baris rusak
df = pd.read_csv('../Dataset/Dataset_Cyber_Security_Indexes_Cleaned.csv', quotechar='\"', on_bad_lines='skip')

# Paksa CEI dan NCSI menjadi numerik, abaikan baris error
df['CEI'] = pd.to_numeric(df['CEI'], errors='coerce')
df['NCSI'] = pd.to_numeric(df['NCSI'], errors='coerce')

# Hapus baris yang memiliki NaN setelah konversi
df = df.dropna(subset=['CEI', 'NCSI'])

# Ambil data numerik yang sudah bersih
x = df['CEI']
y = df['NCSI']

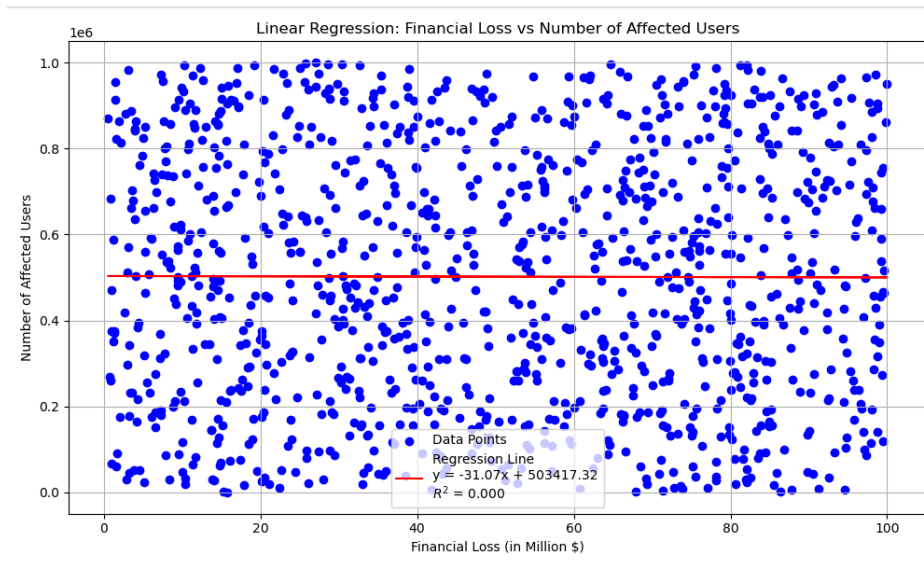
# Linear regression
slope, intercept, r, p, std_err = stats.linregress(x, y)

# Model fungsi
def model(x):
    return slope * x + intercept

y_model = list(map(model, x))

# Plotting
plt.figure(figsize=(10, 6))
plt.scatter(x, y, color='blue', label='Data Points')
plt.plot(x, y_model, color='red', label=f'Regression Line\mathbf{y} = \{slope:.2f\}x + \{intercept:.2f\}\mathbf{R}^2 = \{r**2:.3f\}')
plt.xlabel('CEI (Cybersecurity Exposure Index)')
plt.ylabel('NCSI (National Cyber Security Index)')
plt.title('Linear Regression: CEI vs NCSI')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```

Gambar 9.1.1. Code Python Data Analysis Dataset Cyber Security Indexes



Gambar 9.1.2. Output Code Python Data Analysis Dataset Cyber Security Indexes

9.2 Data Analisis Global Cyber Security Indexes

Pada dataset Global Cyber Security Indexes, variabel yang digunakan meliputi *cyber exposure index* (CEI) sebagai variabel independen dan *national cyber security index* (NCSI) sebagai variabel dependen. Berikut merupakan kode yang digunakan untuk mencari regresi linier, beserta dengan keluaran dari program.

```
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats

# Load CSV dengan penanganan koma dan baris rusak
df = pd.read_csv('../Dataset/Dataset_Cyber_Security_Indexes_Cleaned.csv', quotechar='\"', on_bad_lines='skip')

# Paksa CEI dan NCSI menjadi numerik, abaikan baris error
df['CEI'] = pd.to_numeric(df['CEI'], errors='coerce')
df['NCSI'] = pd.to_numeric(df['NCSI'], errors='coerce')

# Hapus baris yang memiliki NaN setelah konversi
df = df.dropna(subset=['CEI', 'NCSI'])

# Ambil data numerik yang sudah bersih
x = df['CEI']
y = df['NCSI']

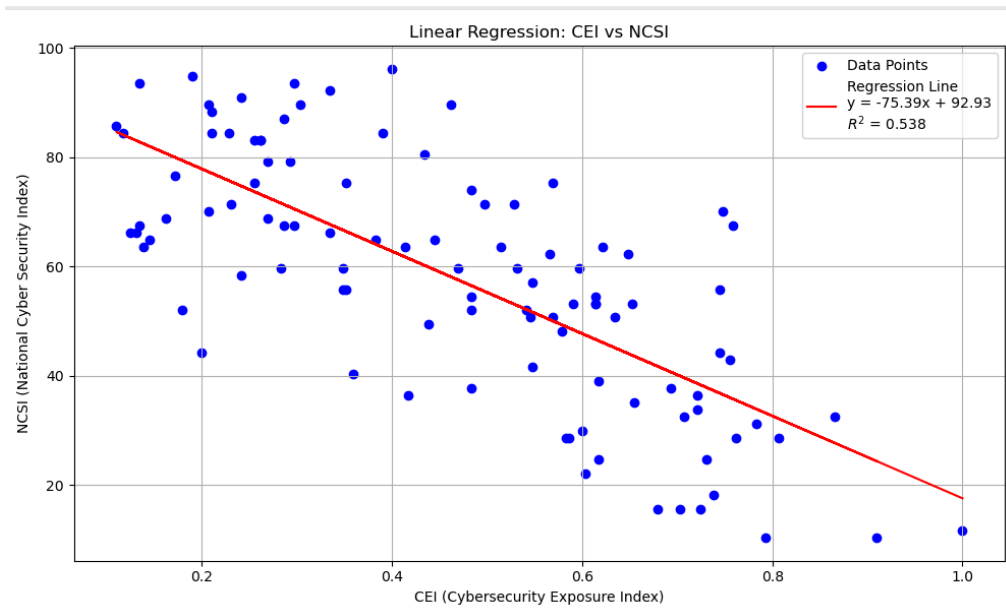
# Linear regression
slope, intercept, r, p, std_err = stats.linregress(x, y)

# Model fungsi
def model(x):
    return slope * x + intercept

y_model = list(map(model, x))

# Plotting
plt.figure(figsize=(10, 6))
plt.scatter(x, y, color='blue', label='Data Points')
plt.plot(x, y_model, color='red', label=f'Regression Line\ny = {slope:.2f}x + {intercept:.2f}\nR^2 = {r**2:.3f}')
plt.xlabel('CEI (Cybersecurity Exposure Index)')
plt.ylabel('NCSI (National Cyber Security Index)')
plt.title('Linear Regression: CEI vs NCSI')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```

Gambar 9.2.1. Code Python Data AnalisisDataset Global Cybersecurity Threats



Gambar 9.2.1. Code Python Data AnalysisDataset Global Cybersecurity Threats

KESIMPULAN

Dalam keamanan siber, terdapat berbagai atribut data yang muncul ketika mengalami penyerangan ancaman siber. Berdasarkan hasil analisis, dapat disimpulkan bahwa tidak terdapat korelasi antara ketiga atribut, yaitu jumlah pengguna yang terdampak, kerugian finansial, dan waktu penyelesaian insiden. Faktor-faktor teknis seperti hal tersebut tidak menentukan dampak yang signifikan dalam suatu serangan siber. Namun, faktor-faktor kontekstual seperti kompleksitas serangan, jenis ancaman, serta bidang yang diserang lebih memengaruhi besarnya dampak suatu serangan siber.

Sementara itu, korelasi antar indeks global keamanan siber seperti Cybersecurity Exposure Index (CEI), National Cyber Security Index (NCSI), Global Cybersecurity Index (GCI), dan Digital Dependency Level (DDL) saling memiliki hubungan yang kuat. Hal tersebut dapat dibuktikan dengan menggunakan fitur spreadsheet *correl(data y: data x)* untuk menghitung koefisien korelasi dan selalu didapatkan koefisien korelasi antara $-1 \leq x \leq -0,5$ atau $0,5 \leq x \leq 1$. Terdapat korelasi negatif antara CEI dengan NCSI dan GCI yang mengindikasikan bahwa semakin tinggi kesiapan suatu negara terhadap keamanan siber, maka tingkat kerentanannya terhadap serangan siber semakin rendah. Selain itu, terdapat korelasi positif antara NCSI dan GCI yang menyimpulkan bahwa negara dengan komitmen tinggi terhadap keamanan siber akan memiliki tingkat persiapan yang tinggi juga. Kemudian indeks DDL memiliki korelasi positif antara NCSI dan GCI, tetapi memiliki korelasi negatif dengan CEI, yang menyimpulkan bahwa semakin tinggi ketergantungan suatu negara terhadap teknologi digital, maka tingkat kesiapan dan kesadaran keamanannya juga semakin tinggi, serta semakin rendah mengalami ancaman siber.

Negara tidak seharusnya hanya fokus pada satu aspek saja namun perlu mempertimbangkan faktor-faktor lain juga. Oleh karena itu, setiap negara harus memperkuat infrastruktur teknologi mereka dengan meningkatkan kesadaran bahwa keamanan siber merupakan isu yang tidak boleh dianggap remeh. Dengan begitu, dunia dapat memiliki ketahanan digital yang kuat demi menghadapi tantangan global mendatang.

LAMPIRAN

ANGGOTA KELOMPOK	PEMBAGIAN KERJA KELOMPOK
Geraldo Artemius	Penyusun PPT, Penyusun Laporan Bab 2, 4, 5, dan Kesimpulan, Presentasi di Video.
Mikhael Adrian Yonatan	Penyusun PPT, Penyusun Laporan Bab 4, 5, dan 6, Presentasi di Video.
Junior Natra Situmorang	Pencari Dataset, Penyusun Laporan Bab 4 dan 7, Presentasi di Video, Coder untuk Data Cleansing.
Reynard Nathanael	Penyusun PPT, Penyusun Laporan Bab 4, 5, dan 8, Presentasi di Video.
Nicholas Luis Chandra	Penyusun PPT, Pencari Dataset, Penyusun Laporan Bab 1, 3 ,dan 9, Editor Video, Presentasi di Video, Coder untuk Data Analysis.

Tabel 11.1 Lampiran Pembagian Kerja Kelompok