Documentation

Here's the description of the columns in the cyber dataset.

Basic information

Column name	unit	Description of the column
company_name	string	Name of the company
date_published	date	Date of announcement
date_of_breach	date	Date of incident
date_of_breach_end	date	Date of end of incident
total_records_breached	integer	Total number of breached records
pii_breached	boolean	If Personal Identifiable Information (PII) was breached
num_pii_breached	integer	Number of PII records breached
ssn_breached	boolean	If Social Security Numbers (SSN) was breached
num_ssn_breached	integer	Number of SSN records breached
phi_breached	boolean	If Protected Health Information (PHI) was breached
num_phi_breached	integer	Number of PHI records breached
pci_breached	boolean	If Payment Card Information (PCI) was breached
num_pci_breached	integer	Number of PCI records breached

`pfi_breached`	boolean	If Personal Financial Information (PII) was breached
`num_pfi_breached`	integer	Number of PFI records breached
`passwords_breached`	boolean	If Passwords were breached
`num_passwords_breached`	integer	Number of Passwords breached

CIS controls affected in the incident

Column name	unit	Description of the column
`cis_1_affected`	boolean	CIS control #1 affected Inventory and Control of Hardware Assets
`cis_2_affected`	boolean	CIS control #2 affected Inventory and Control of Software Assets
`cis_3_affected`	boolean	CIS control #3 affected Continuous Vulnerability Management
`cis_4_affected`	boolean	CIS control #4 affected Controlled Use of Administrative Privileges
`cis_5_affected`	boolean	CIS control #5 affected Secure Configuration for Hardware and Software on Mobile Devices, Laptops, Workstations and Servers
`cis_6_affected`	boolean	CIS control #6 affected Maintenance, Monitoring and Analysis of Audit Logs
`cis_7_affected`	boolean	CIS control #7 affected Email and Web Browser Protections
`cis_8_affected`	boolean	CIS control #8 affected Malware Defenses
`cis_9_affected`	boolean	CIS control #9 affected

`cis_9_affected`	boolean	Limitation and Control of Network Ports, Protocols and Services
`cis_10_affected`	boolean	CIS control #10 affected Data Recovery Capabilities
`cis_11_affected`	boolean	CIS control #11 affected Secure Configuration for Network Devices, such as Firewalls, Routers and Switches
`cis_12_affected`	boolean	CIS control #12 affected Boundary Defense
`cis_13_affected`	boolean	CIS control #13 affected Data Protection
`cis_14_affected`	boolean	CIS control #14 affected Controlled Access Based on the Need to Know
`cis_15_affected`	boolean	CIS control #15 affected Wireless Access Control
`cis_16_affected`	boolean	CIS control #16 affected Account Monitoring and Control
`cis_17_affected`	boolean	CIS control #17 affected Implement a Security Awareness and Training Program
`cis_18_affected`	boolean	CIS control #18 affected Application Software Security
`cis_19_affected`	boolean	CIS control #19 affected Incident Response and Management
`cis_20_affected`	boolean	CIS control #20 affected Penetration Tests and Red Team Exercises
`originated_from_3rd_party`	boolean	If the incident originated from a third party

Information on losses occurred as result of the incident

Column name	unit	Description of the column
total_loss	dollars	Total loss (in $)
loss_total_sources	string array	Sources for total loss
loss_dte_cost	dollars	Losses concerning Determination of Cyber Event Extent (in $)
loss_dte_sources	string array	Sources concerning Determination of Cyber Event Extent
loss_ruf_cost	dollars	Losses concerning Replacement or Upgrade of Faulty Controls (in $)
loss_ruf_sources	string array	Sources concerning Replacement or Upgrade of Faulty Controls
loss_rvl_cost	dollars	Losses concerning Review of Necessary Legal and Regulatory Actions (in $)
loss_rvl_sources	string array	Sources concerning Review of Necessary Legal and Regulatory Actions
loss_nmr_cost	dollars	Losses concerning Notification, Credit Monitoring, Credit Restoration (in $)
loss_nmr_sources	string array	Sources concerning Notification, Credit Monitoring, Credit Restoration
loss_vrf_cost	dollars	Losses concerning Vendor and Regulatory Fines and Compliance (in $)
loss_vrf_sources	string array	Sources concerning Vendor and Regulatory Fines and Compliance
loss_trf_cost	dollars	Losses concerning Third Party Litigation and Damages (in $)
loss_trf_sources	string array	Sources concerning Third Party Litigation and Damages

Column name	unit	Description of the column
`link_sources`	string array	Additional sources of information

Financial information of the company at the time of the announcement of the incident

Column name	unit	Description of the column
`stock_symbol`	string	Stock symbol
`stock_date`	date	Date of the stock price before the incident (max 30 days)
`adjusted_close_price`	dollars	Adjusted close price of the stock
`num_employees`	integer	Number of employees
`industrial_sector`	string	Industry Classification Benchmark (ICB) sector

Each of the following financial features is available at different time intervals relative to the date of announcement. For each of those features `<feat>`, we have available the values for:

- `<feat>_m1` : The `<feat>` value on the fiscal quarter previous to the date of announcement
- `<feat>_p1` : The `<feat>` value on the fiscal quarter after the date of announcement
- `<feat>_p2` : The `<feat>` value 2 fiscal quarters after the date of announcement
- `<feat>_p3` : The `<feat>` value 3 fiscal quarters after the date of announcement
- `<feat>_p4` : The `<feat>` value 4 fiscal quarters after the date of announcement

Column name	unit	Description of the column
`statement_date_*`	date	Publication date of the quarterly financial data
`revenue_*`	dollars	Revenue
`nop_inc_*`	dollars	Non-Operating Income- Total
`sga_exp_*`	dollars	Selling, General and Administrative Expenses
`rd_exp_*`	dollars	Research and Development Expense
`op_exp_*`	dollars	Operating Expense- Total

eps_op_*	dollars	Earnings Per Share from Operations
ir_exp_*	dollars	Interest and Related Expense- Total
pretax_inc_*	dollars	Pretax Income
inc_tax_*	dollars	Income Taxes - Total
disc_op_*	dollars	Discontinued Operations
net_inc_*	dollars	Net Income
ycc_equi_*	dollars	Yearly Cash and Cash Equivalents - Increase (Decrease),
st_inv_*	dollars	Short-Term Investments - Total
receiv_*	dollars	Receivables - Total
invent_*	dollars	Inventories - Total
cur_assets_*	dollars	Current Assets, Total
cur_assets_oth_*	dollars	Current Assets, Other, Total
lt_inv_*	dollars	Total Long-term Investments
ppe_*	dollars	Property, Plant and Equipment - Total (Gross)
goodwill_*	dollars	Goodwill (net)
int_assets_*	dollars	Intangible Assets - Total
dda_*	dollars	Depreciation, Depletion and Amortization (Accumulated)
assets_*	dollars	Assets - Total
debt_liab_*	dollars	Debt in Current Liabilities
curr_liab_*	dollars	Current Liabilities - Total
lt_liab_*	dollars	Long-Term Liabilities (Total)
lt_debt_*	dollars	Long-Term Debt - Total
liab_oth_*	dollars	Liabilities - Other

`liab_*`	dollars	Liabilities - Total
`pp_stock_redeem_*`	dollars	Preferred/Preference Stock - Redeemable
`pp_stock_cap_*`	dollars	Preferred/Preference Stock (Capital) - Total
`ret_earn_*`	dollars	Retained Earnings
`tr_stock_*`	dollars	Treasury Stock - Total (All Capital)
`cap_ssp_res_*`	dollars	Capital Surplus/Share Premium Reserve
`sth_eq_*`	dollars	Stockholders Equity - Total
`dep_amr_*`	dollars	Depreciation and Amortization - Total