Feature Name	Feature Description	Justification	Mean	Median	Standard Deviation	Feature Name	Mean (Breached)	Median (Breached)	Standard Deviation (Breached)	Feature Name	Mean (Not Breached)	Median (Not Breached)	Standard Deviation (Not Breached)
employees	number of employees	With more employees, a company has more opportunities for human error that could lead to a breach	18.451	2.8	67.897	employees	73.282	27.826	98.072	employees	17.515	2.7	66.883
SIC4	The industry indicator	breached	4835.509	4923	1896.967	SIC4	5502.377	5912	1543.896	SIC4	4823.832	4922	1900.515
Breach	Whether a company was breached that year	This is our outcome variable that we are trying to predict.	0.019	0	0.138	Breach	1	1	1	Breach	0	0	0
BreachType	experienced	Outcome variable we are trying to predict	n/a	n/a	n/a	BreachType	n/a	n/a	n/a	BreachType	n/a	n/a	n/a
stockholderequity	Accounting measure, The total quity held by shareholders	Companies who have a higher sharholder equity, are likely to have stronger finances making them more suseptable to breaches	2951.532	486.926	10526.637	stockholderequity	14345.661	2948.915	28755.338	stockholderequity	2751.246	474.311	9796.411
retained earnings	Accounting measure, The profits that a company has held for reinvestment	A company that has higher retained earnings and financial stability could be more at risk for a breach, but could also indicate better financials to spend on preventing data breaches	1649.594	152.928	9439.7895	retained earnings	8620.469	2034.9	20691.833	retained earnings	1526.383	147.601	9070.765
netincome	Accounting measure, Total profit of a company after taxes and costs	threat actors would like to breach for financial gain, however this also might indicate better cybersecurity investments	367.432	40.057	2301.4089	netincome	1685.674	468.583	4448.66	netincome	344.245	38.975	2238.767
cash	Accounting measure, total cash held by a company	The more cash that a company has, the more they may be able to invest in increasing their cybersecurity and reducing risks of breaches	670.182	74.242	3340.63	cash	3047.308	732.612	6546.494	cash	628.279	72.32	3240.88
liabilities	Accounting measure, total obligations that a company owes	A company with higher liabilities could have less ability to invest in increaing cybersecurity and therefore be at more risk for cyber breaches	12040.77	860.159	84672.481	liabilities	93128.467	6239.171	316182.82	liabilities	10611.426	835.345	73647.632
TOT_C	Total firm social responsibility concerns	Companies that have higher social responsibility concerns may be more suseptible to cyber breaches due to records of poor social responsibility	1.875	2	1.783	TOT_C	3.1625	2	2.725	TOT_C	1.845	2	1.744
TOT_S	Total firm social responsibility strengths	Companies that have higher social responsibility streangth are less likely to be suseptible to cyber breaches due to records of proactive risk management, and increased cybersecurity	1.566	1	2.594	TOT_S	3.725	2	4.376	TOT_S	1.516	1	2.516
tobinq	Accounting measure, captures IT spending	show that companies investments in their technology and cybersecurity which would lower its risk of breach, however lower IT spending could make the company vulnerable	1.589	1.12	2.913	tobinq	1.436	1.072	1.44	tobinq	1.591	1.121	2.933
newshits	The number of times a firm is mentioned in the news	When a company is mentioned more times in the news, they can look more attractive to threat actors and therefore be at more risk of a cyber breach	603.02	142	2678.146	newshits	3270.299	663.5	7187.44	newshits	555.195	140	2499.839
IsHealthCare	Is company in the healthcare industry	Handles highligh sensitive data, making it a target for breaches	0.012	0	0.11	IsHealthCare	0.013	0	0.111	IsHealthCare	0.012	0	0.11
IsRetail	Is company in the Retail Industry	Deals with a signifiant amount of customer data such as payment and personal information, making it a high risk target	0.002	0	0.044	IsRetail	0.008	0	0.091	IsRetail	0.002	0	0.043
IsIT	Is company in the IT industry	IT industry may be at target because of the their area of expertise	0	0	0	IsIT	0	0	0	IsIT	0	0	0
IsCommunicationServices	Is company in the Communications industry	this may be an attrative target if the attach aims to intercept or disrupt communications that damage operations	0.025	0	0.155	IsCommunicationServices	0.05	0	0.218	IsCommunicationServices	0.024	0	0.153
IsFinancialServices	Is company in the Financial industry	Due to the senestive financial data they handle, they're seen as high risk due to their value	0.228	0	0.419	IsFinancialServices	0.246	0	0.431	IsFinancialServices	0.227	0	0.419
IsEntertainment	Is company in the Entertainment industry	Companies in the entertainment industry could be more suseptible to cyber breaches due to their content they produce and the personal data that they hold, which could make them more attractive to threat actors	0.008	0	0.09	IsEntertainment	0	0	0	IsEntertainment	0.008	0	0.091
ROE	Return on Equity	A lower ROE may indicate poorer preformance regarding a company's resources, making it a more attractive target. While a higher ROE may indicate a company's successful performance, signalling to a lack of vulnerabilities	-0.013	0.102	4.388	ROE	0.154	0.144	1.632	ROE	-0.016	0.101	4.421
NetincomeToLiabilities	Ratio of company's net income to its total liabilities	Higher value would indicate there is more income relative to liabilities, which reveals healthy finances. This may lead to the company investing more on security measures.	0.027	0.063	1.598	NetincomeToLiabilities	0.131	0.074	0.26	NetincomeToLiabilities	0.026	0.062	1.612
TOT_CtoTOT_S	Ratio of company's TOT_C to TOT_S	A higher ratio may indicate that a company has more concerns than strengths, which could negatively impact their reputation with consumers, or the general public. Likely leading to less attraction, less revenue and inability to invest in cybersecurity	1.113	1	1.117	TOT_CtoTOT_S	1.256	1	1.273	TOT_CtoTOT_S	1.109	1	1.112
IsGoverment	Is company/organization a goverment entity	Goverment organizations deal with highly sensetive information of critital infrastructure and personnels. If politically motived, a goverment organization may be at high risk depending on the information they handle	0	0	0	IsGoverment	0	0	0	IsGoverment	0	0	0
IsEducation	Is company/organization tied to educational institutions	Educational institutions tend to have weaker security infrastructures, considering that aspect with the large personal data and information they possess, they may be high risk targets	0.004	0	0.06	IsEducation	0	0	0	IsEducation	0.004	0	0.06
IsManufacturing	Is company in the Manufacturing industry	Breach attachs against industrial control systems (ICS) and operational technology (OT) are becoming more frequent in the manufacturing sector, therefore a company being in the manufacturing industry would increase its chances of being a target	0.36	0	0.48	IsManufacturing	0.192	0	0.394	IsManufacturing	0.363	0	0.481
BreachOccurenceRate	Rate of breach frequency over the years within company	Companies with a higher breach rate may increase the risk of future attacks if initial vulnerabilities were not addressed efficiently and effectively	0.019	0	0.068	BreachOccurenceRate	0.267	0.167	0.168	BreachOccurenceRate	0.014	0	0.054
TotalAssets	Company's total assets	Larger assets number may draw cybercriminals who want to steal or exploit valuable data or interfere with daily operations	14544.97	1389.665	91124.316	TotalAssets	93144.244	7529.1979	318160.62	TotalAssets	12991.628	1349.893	79692.517
ROA	Total return on assets (ROA)	Low ROA could indicate a company has low profitability therefore underinvests in security, leading to higher changes to experience a breach	0.001	0.037	0.906	ROA	0.047	0.047	0.095	ROA	9.80E-05	0.036	0.914
AssetTurnover	Company's asset turnover rate	A high asset turnover rate may indicate the company is effectively utilizing it's assets and aproachs the management of areas such as security seriously, lowering its chances for a breach	0.001	0.037	0.906	AssetTurnover	0.047	0.047	0.095	AssetTurnover	9.80E-05	0.036	0.914
EmployeeToCapEx	Employee investment ratio	A lack of investment in employees may leave gaps in areas such as training and increasing human error, which may increase risk of breach	0.216	0.075	1.113	EmployeeToCapEx	0.137	0.077	0.245	EmployeeToCapEx	0.218	0.075	1.122
CapExToTotalAsset	How much company is investing its assets on CapEx	in secure infrusture and technology, including cybersecurity systems making the company an attractive target	0.033	0.014	0.057	CapExToTotalAsset	0.243	0.119	0.437	CapExToTotalAsset	0.033	0.014	0.057
CashToLiabilities	Cash to liabilities ratio to calculate liabilities	Lower ratio may suggests a company is facing financial instability, which may result in a lack of cybersecurity investing and maintance, increasing the risk of a breach	0.433	0.105	1.485	CashToLiabilities	0.243	0.119	0.437	CashToLiabilities	0.436	0.105	1.497
CashFlowMargin	Conversion ratio of sales to cash	Lower ratio suggests the company's poor financial health, which will likely result in a lack of cybersecurity investing, increasing the risk of a breach	1.912	0.544	89.655	CashFlowMargin	2.801	0.657	20.67	CashFlowMargin	1.896	0.541	90.4
StockholderEquityToLiability	Ratio of company's Stockholder equity to liability	mean fewer resources are available for cybersecurity, making breaches more likely. However, it may also decrease the risk of attacks considering their poor financial health	1.888	0.765	13.611	StockholderEquityToLiability	0.956	0.609	1.476	StockholderEquityToLiability	1.905	0.771	13.729
BreachtoCapEx	Breach to CapEx ratio	the company is still experiencing breaching which suggests gaps in their security practices, systems and measures.	0.003	0	0.062	BreachtoCapEx	0.022	0.002	0.058	BreachtoCapEx	0.003	0	0.062
BreachTotobinq	Breach to IT spending ratio	High ratio indicates that despite high spending on IT, the company is still experiencing breaching which suggests gaps in their security practices, systems and measures.	0.021	0	0.126	BreachTotobinq	0.274	0.161	0.372	BreachTotobinq	0.017	0	0.111
tobinqToEmployee	IT spending per employee ratio	Lower ratio may indicate a lack of IT spending per employee in terms of security training, technology requirements, etc. This may increase a company's	19.351	0.323	454.376	tobinqToEmployee	0.193	0.033	0.556	tobinqToEmployee	19.68	0.333	458.263
EmployeeGrowthRate	Employee growth rate compared to previous year	A high growth rate, especially in a small time frame, leaves more room for human error if company does not consider and implement cybersecurity traninig and processes effictively during onboarding	4.687	-0.043	211.26	EmployeeGrowthRate	-0.238	-0.041	0.581	EmployeeGrowthRate	4.781	-0.043	213.27
NewsHitsImpact	News hit rate changes	An increase in news hits after a breach may attrack future attackers due to reputational damage	1.857	0	33.121	NewsHitsImpact	0.278	0	2.22	NewsHitsImpact	1.89	0	33.458
CapExToTobinq	CapEx to Tobinq investment ratio	Reveals how much of Capital Expenditure spending and IT compares, indicating if enough funding it going towards areas such as privacy, security, etc. measures	315.582	7.569	2612.75	CapExToTobinq	2201.146	51.563	14461.149	CapExToTobinq	282.53	7.378	1798.518

OLD

Clients	Total number of clients	More clients may increase the interaction points of an attack, through client related vulnerabilities
Contractors	Total number of contractors	With contractors often having acess to company systems, there may be a risk if they're not properly trained
GlobalPrescense	Total of international offices	By operating in different regions, there may be different security measures, regulations, and an increase in potential attackers
SecurityBudget	Annual budget for security measures	Size of a company's security budget may indicate how robust and effictive their cybcersecuirty measures are. Often, a higher security budget indicates more resources and stronger cybersecurity practies within the company as it's invested in training, incident response, etc.
SecurityScore	Organization's risk and cybersecurity posture metric	A company's overall risk and cybcersecurity posture score may indiciate its strengths and weaknesses, a lower score making it more vulnerable to attacks
SecurityTraninigFrequency	Frequency of employee security training	Higher frequence of training can reduce human error, which often leads to the cause of the breach
EmployeeSecurityTraninig	Percentage of employees who completed security training	
ThirdPartyVendors	Total number of third party vendors	
ThirdPartyVendorRisk	Risk level of third party vendors	
CybersecurityMaturityLevel	Maturity number of overall cyber practices	
MFAUsage	Total of employee accounts using Multi-Factor Authentication	
ITStaff	Total of IT Staff, ratio to total employees	
IncidentResponse	Time taken to response to security indicents	
PhishingTestingSucess	Percentage of postive employee preformance during phishing testing	
PatchFrequency	Frequency at which security patches are applied to system(s)	
DataSensetivityScore	Sensetivity score for type of data the company handles	
CybersecurityCoverage	Percentage of systems or applicatoins within the company that are protected by security application or infrastructure	
MarketShare	Percentage of markeshare within industry	