Abstract:

The digital transformation of the healthcare industry has made treatment more accessible, but it has also become a victim of data breaches. This study analyzes healthcare data breaches and aims to improve confidentiality. Hacking/IT incidents are the most common form of attack, followed by unauthorized internal disclosures. The frequency, magnitude, and financial losses due to breaches are increasing. The study uses the simple moving average method and the simple exponential soothing method to examine data breaches and their costs, providing more reliable forecasting results.

Intro and some statistics:

From 2005 to 2019, 249.09 million individuals were affected by healthcare data breaches, with 157.40 million affected in the last five years. In 2018, the healthcare industry faced 536 breaches, with 41.2 million records exposed or stolen. The average cost of a data breach in 2019 was $3.92 million, the highest in the USA. The average cost of a breach increased by 12% from 2014 to 2019, with the cost of a breached record in the healthcare sector rising 19.4%.

Methodology:

The data analysis method involves compiling data, applying sum, percentage, and average methods, extracting patterns, understanding sources and consequences, and using time series analysis for healthcare data breach forecasting.

Data source:

Data was sourced from the PRC Database, the HIPAA Journal, OCR Reports, Ponemon Institute Reports, and Verizon-DBIR. The PRC Database provides a comprehensive record of data breaches, with over 10 billion user records compromised since 2005. The HIPAA Journal provides compliance guidelines, while Ponemon Institute Reports and Verizon-DBIR provide data breach costs and invasion instances globally.

Result:

From 2005 to 2019, over 10 billion records were exposed in various sectors, with the healthcare sector being the most affected. Hacking (HACK) was the primary cause of data breaches, with over 64% of health data breached between 2005 and 2019. In the last five years, hacking incidents exposed more than 92% of records. Other types of attacks included PHYS and PORT, causing 14.39% and 9.51% of the total exposed records, respectively. HIPAA and OCR reports revealed that hacking/IT incidents were the main cause behind healthcare data breaches, affecting 255.18 million people from 3051 incidents from 2010 to 2019.  
  
The main types of attacks used to breach protected health data include hacking/IT incidents, unauthorized access/internal disclosure, theft/loss, and improper disposal. However, theft/loss and improper disposal have shown a decreasing trend, while hacking/IT incidents and unauthorized internal disclosures have increased significantly. Hacking/IT incidents increased by 73.4% in 2019 from 2018, while unauthorized internal disclosure, theft/loss, and improper disposal decreased by 0.7%, 7.8%, and 22.22%, respectively, from 2018 to 2019.  
  
The average cost of data breach in the healthcare industry is $6.45 million, up from $3.92 million in 2019. The cost of each breached record in the sector increased by 5.14% in 2019. The SMA and SES methods of time series analysis were used for healthcare data breach and cost forecasting, with SMA providing more accurate forecast results than SES.

[Paper link](https://www.mdpi.com/2227-9032/8/2/133)