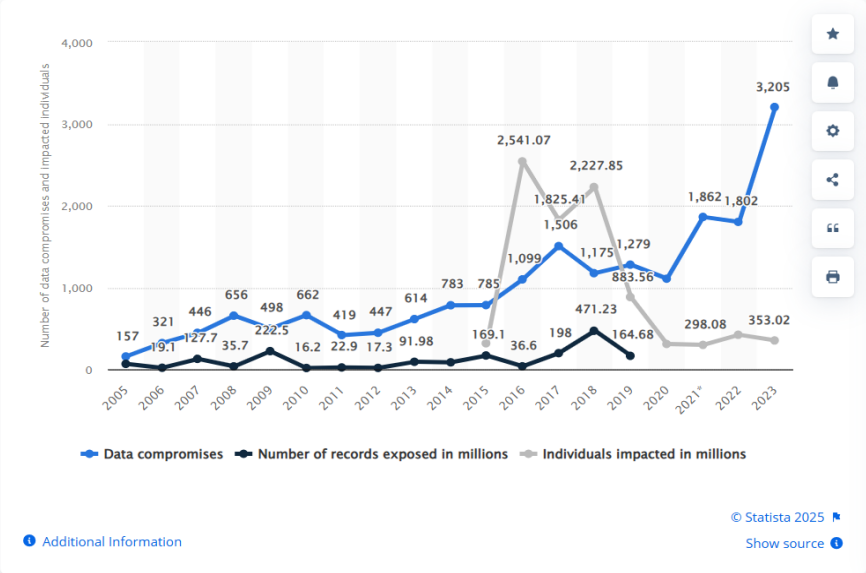
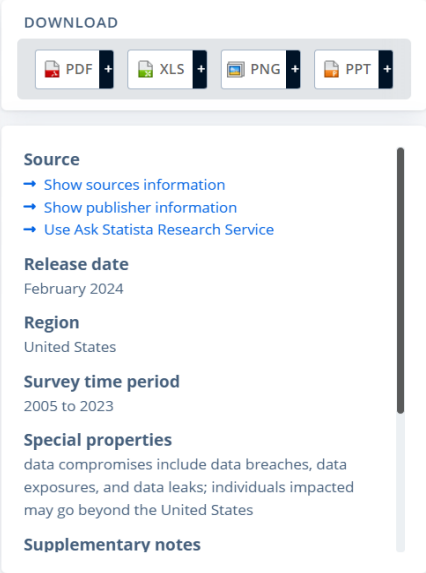
Annual number of data compromises and individuals impacted in the United States from 2005 to 2023





Keyways AI can contribute to data breaches in finance:

- **Increased Attack Surface:**

  Integrating AI into financial operations expands the potential entry points for attackers, as they can target vulnerabilities within the AI models themselves, data pipelines, or the infrastructure supporting them.
- **Data Leakage through Training Data:**

  The vast amount of sensitive financial data used to train AI models can inadvertently leak information if not properly anonymized or secured, exposing customer details to potential attackers.
- **Model Manipulation:**

  Cybercriminals can manipulate AI models through "adversarial attacks" by feeding them carefully crafted data to produce incorrect outputs or reveal sensitive information.
- **Lack of Explainability ("Black Box" Problem):**

  The complex nature of many AI algorithms makes it difficult to understand how they reach decisions, hindering the identification and mitigation of security vulnerabilities.
- **Automated Attacks:**

  AI can be used to automate malicious activities like phishing campaigns, credential stuffing, or even complex financial fraud schemes, making attacks more efficient and harder to detect.
- **Insider Threats:**
  Malicious insiders with access to AI systems could exploit their knowledge to manipulate data or launch targeted attacks.

Examples of AI-related data breaches in finance:

- **AI-powered Fraud Detection Bypass:**

  Cybercriminals can use AI to modify transaction data in a way that bypasses traditional fraud detection systems powered by AI.

- **AI Chatbot Exploits:**

  Hackers can exploit vulnerabilities in AI-powered customer service chatbots to gain access to sensitive customer information.

- **Data Breach through Model Reverse Engineering:**
  Advanced attackers may attempt to reverse engineer an AI model to extract sensitive information embedded within its training data.

Mitigating the risks:

- **Robust Data Security Practices:**

  Implementing strong data encryption, anonymization techniques, and access controls for training data.

- **Model Monitoring and Validation:**

  Continuously monitoring AI models for anomalies and potential vulnerabilities, including regular testing and validation processes.

- **Explainable AI (XAI):**

  Utilizing techniques that provide insights into how AI models make decisions to identify potential biases and security risks.

- **Cybersecurity Expertise:**
  Building a team with specialized knowledge in AI security to proactively identify and address emerging threats.