BEHIND THE BAIT:

Patterns in Phishing Data

https://github.com/evageck/phishing-intelligence-report

Eva Geck

I used real phishing data and simulated training results to identify high-risk patterns and recommend targeted security improvements.

Associate Data Informatics Analyst at ServiceNow

- Analyze with SQL

Basic knowledge of SQL – You don't need to be an expert, but understanding how to query data will help you in this role.

- Build Dashboards

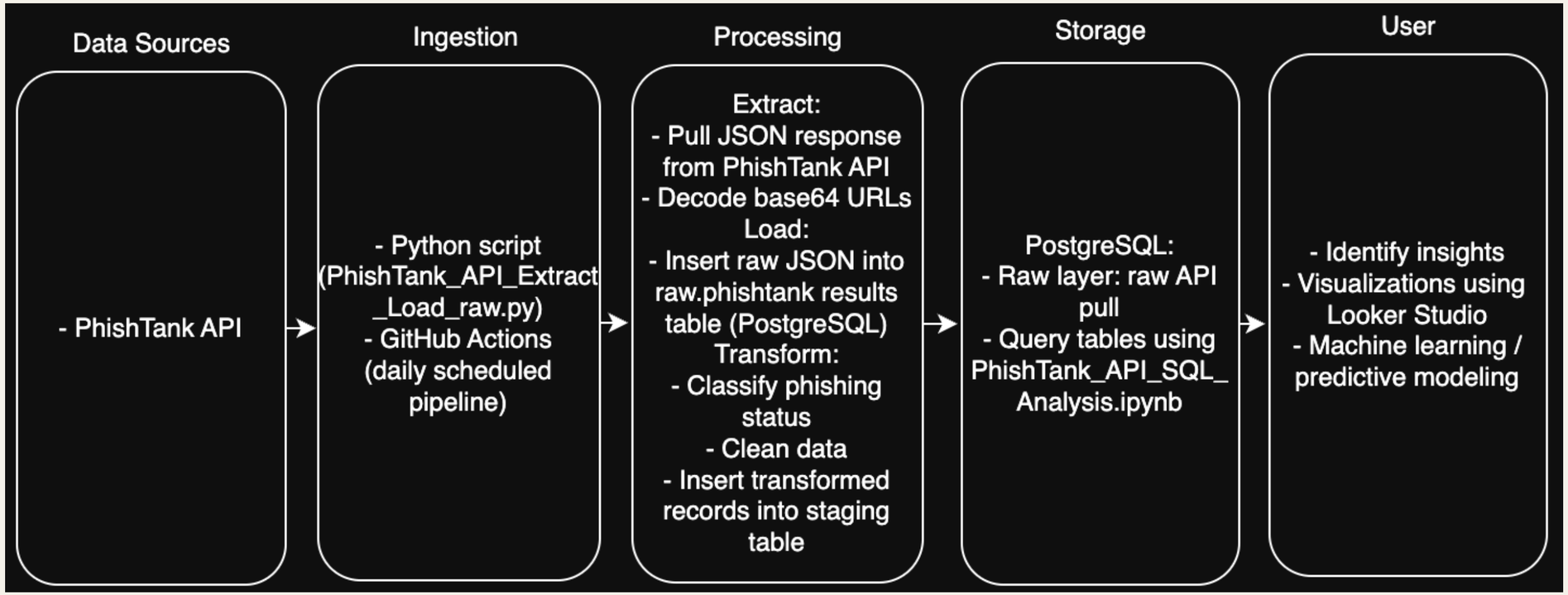
Present Insights Effectively: Communicate findings clearly to both technical and non-technical audiences, helping drive data-informed decisions.

- Use Data Warehouses

Work with Data Warehouses: Learn how to connect and analyze data from Snowflake and other data sources to support business needs.

PhishTank API Integration

https://phishtank.org/api_info.php

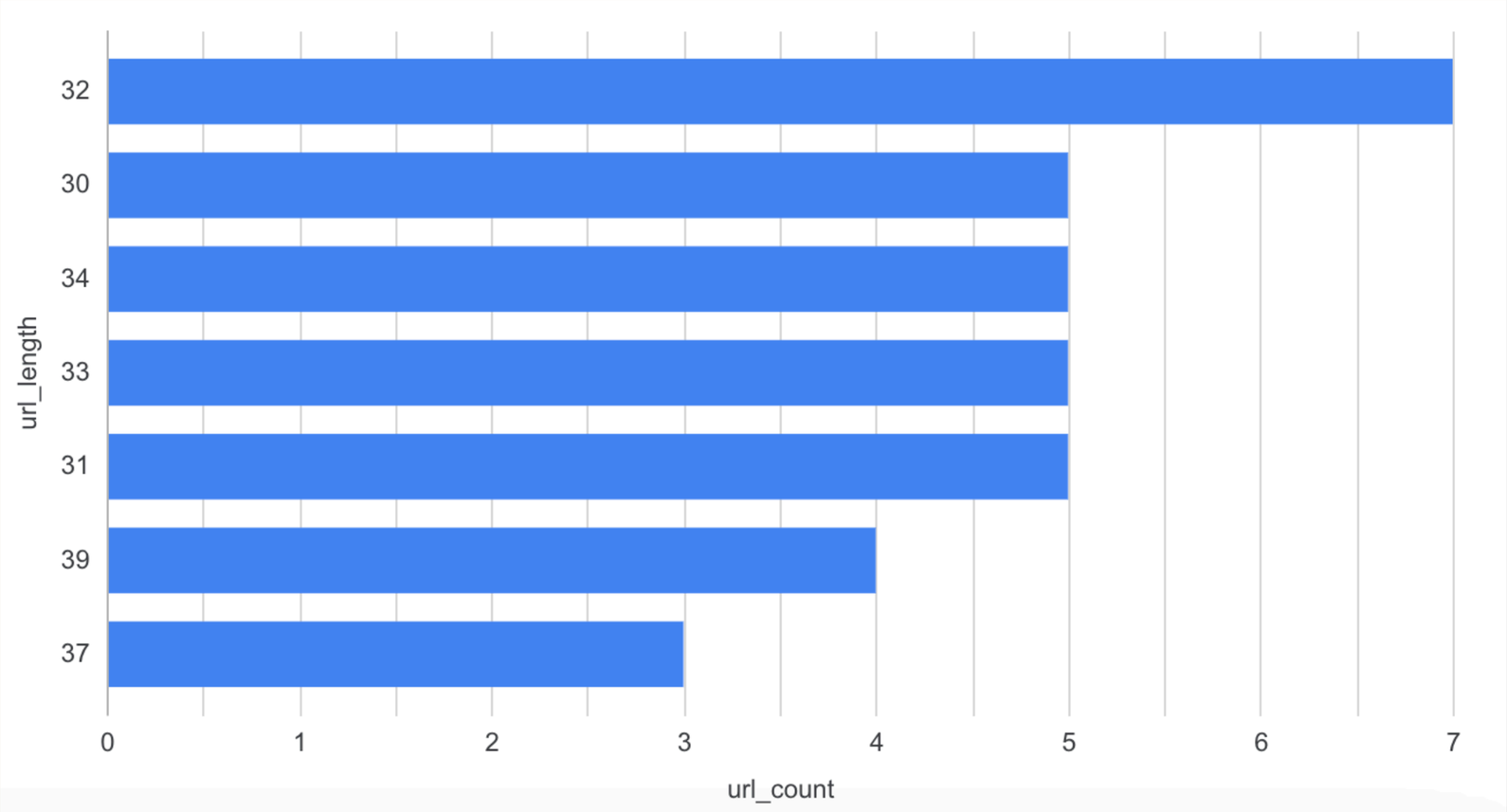


Key Metrics: URL, HTTPS, IP, Length

Job Use: Detect phishing trends to inform defense

Most phishing URLs are 32 characters

What are the most common URL lengths among verified phishing websites?

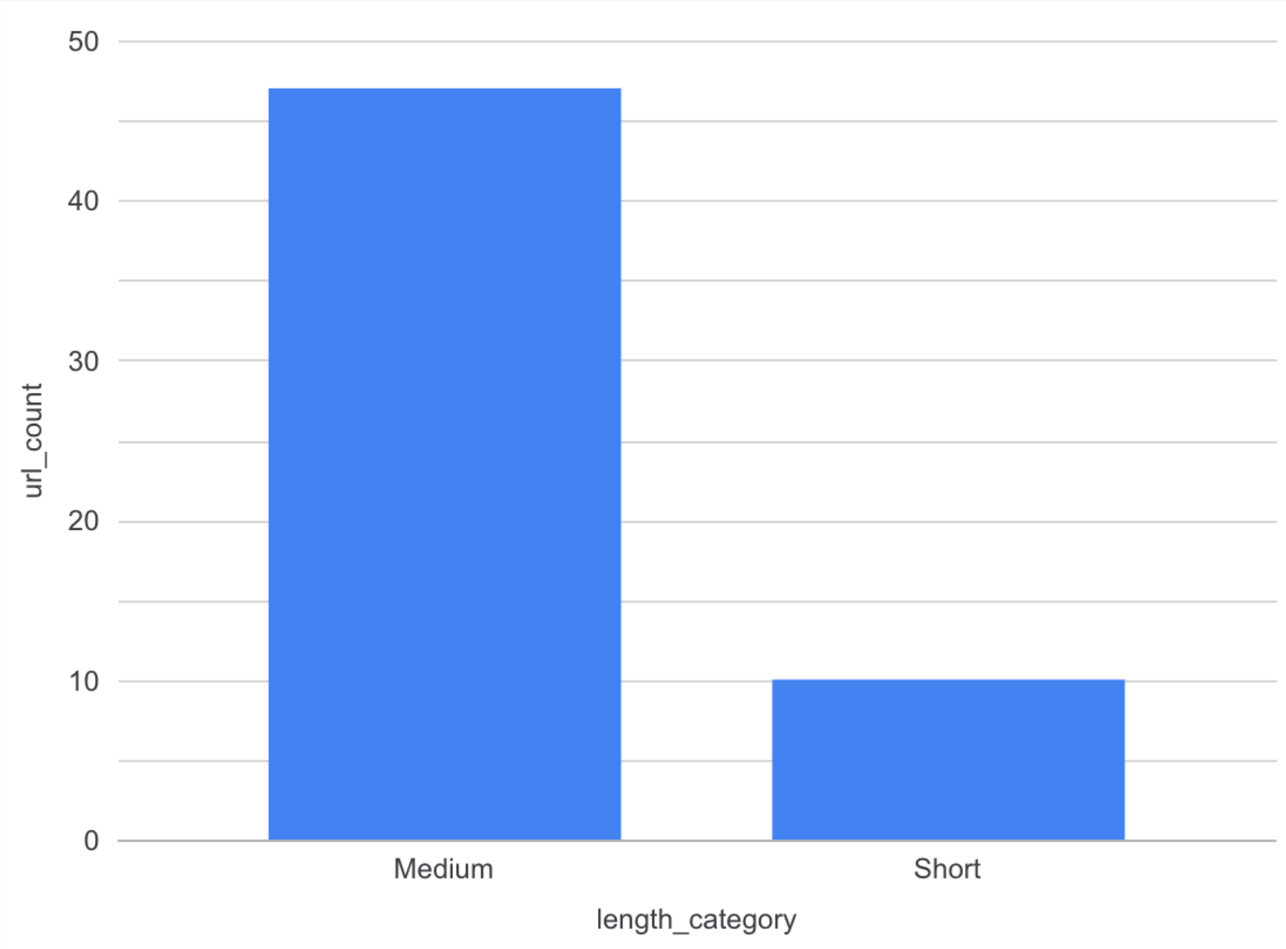


Recommendation: Flag 32-character URLs if other suspicious signs (like no HTTPS) are present

Prediction: Filtering 32-character URLs with other red flags will improve early detection and reduce user risk

Medium-length URLs dominate phishing links

What is the most frequently observed URL length category—Short, Medium, or Long—among verified phishing URLs?

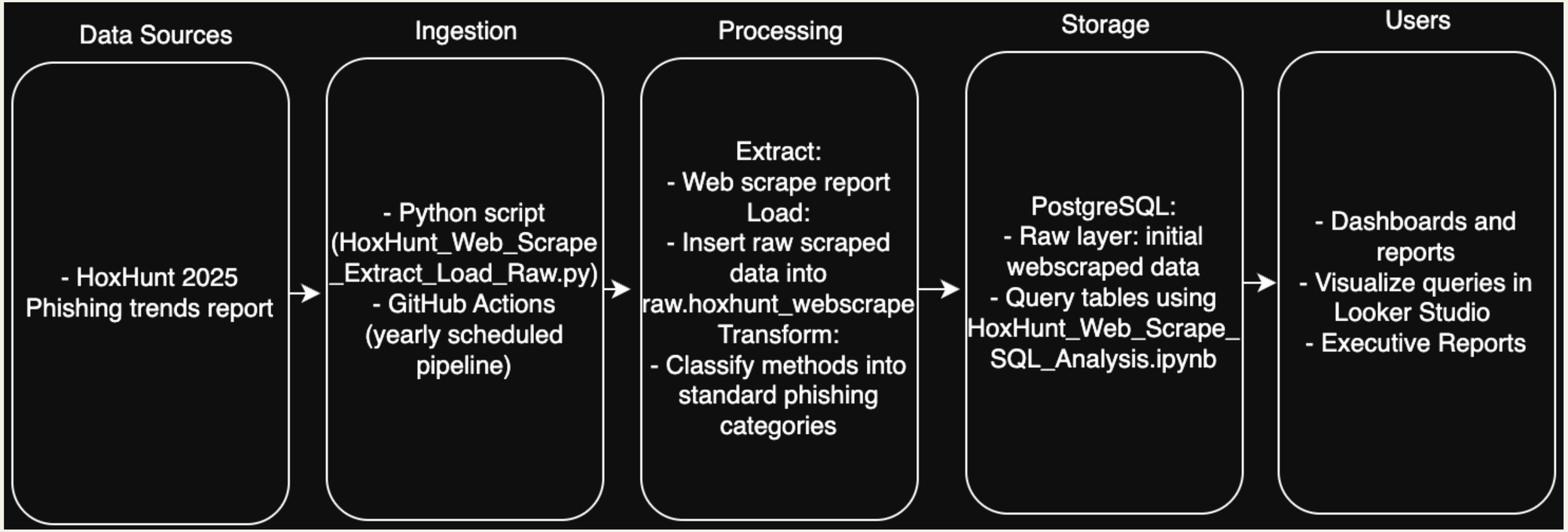


Recommendation: Incorporate URL length category as a feature in phishing detection models

Prediction: Enhancing models with categorical length data will improve prediction accuracy

Web-Scraped Hoxhunt Training Data

https://hoxhunt.com/guide/phishing-trends-report#key-phishing-statistics-for-2025

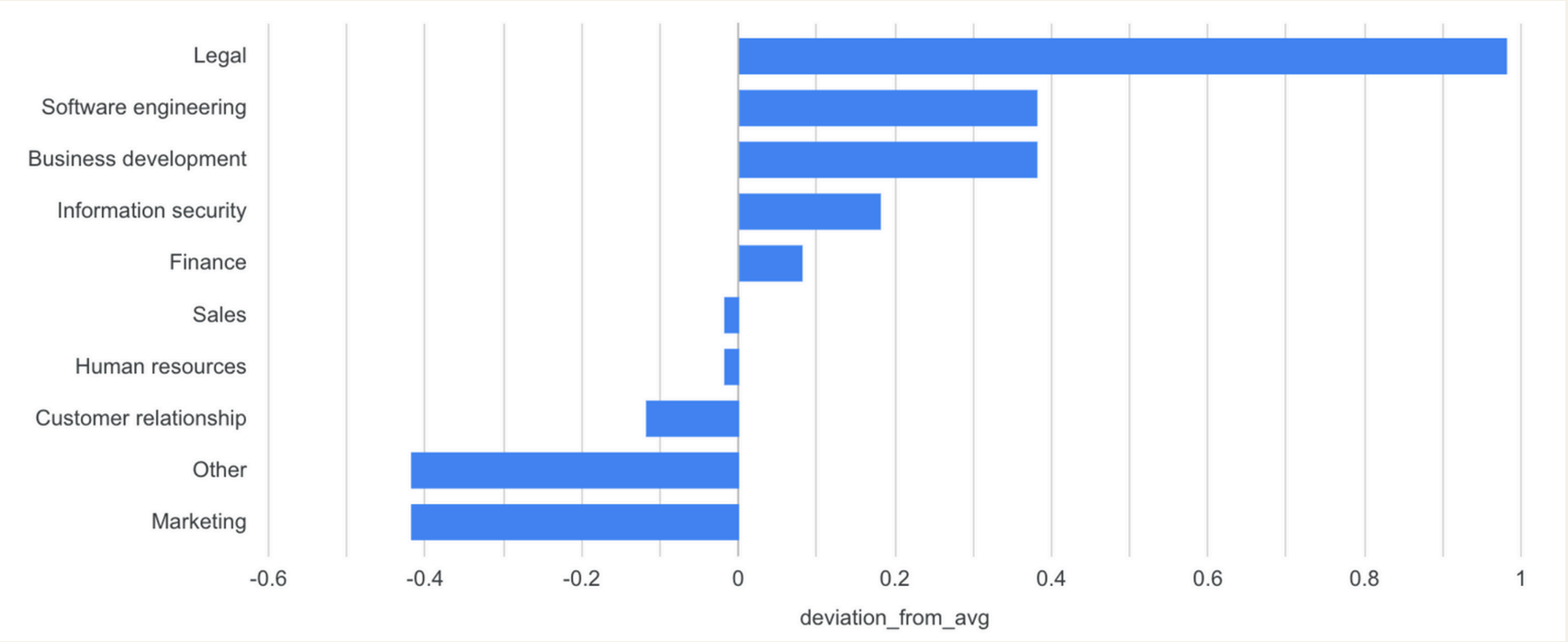


Key Metrics: Industry, department, failure rates

Job Use: Evaluate security training effectiveness

Some Departments Fail More Than Others

Which departments are performing significantly above or below the organization-wide average in phishing training failure rate?

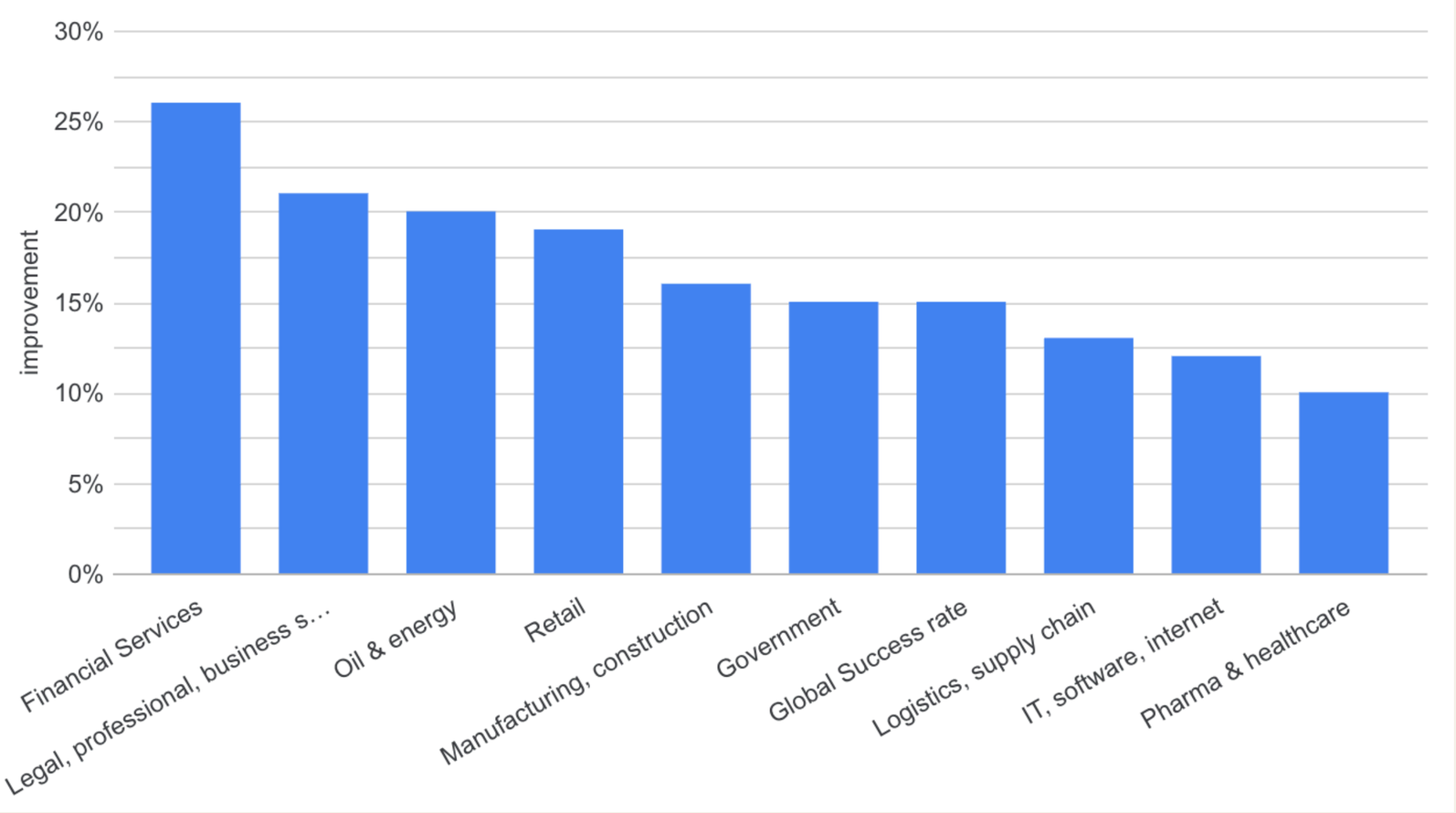


Recommendation: Prioritize targeted training for high-risk departments like Legal, Software Engineering, and Business Development

Prediction: Focused interventions will reduce failure rates

Some Industries Lag in Training Gains

Which industries demonstrate the greatest improvement in phishing training success from month 0 to month 12, and how do they rank?



Recommendation: Enhance training for lagging sectors like Pharma & Healthcare and IT to close performance gaps

Prediction: Improved support will reduce risk disparities and elevate overall phishing resilience

Why I’m Ready for This Role

- SQL and visualization skills match job scope
- Analyzed phishing URLs and training risks
- Built automated workflows (API + scrape)
- Delivered actionable insights via dashboards
- Aligns with ServiceNow’s mission and impact