M3 Project Deliverable 3

Information Visualization Report:

Common software weaknesses in CISAs Known Exploitable Catalog

**Project Background:**

In my project, I analyze data from three major sources in the cybersecurity field: the Cybersecurity and Infrastructure Security Agency (CISA), the National Institute of Standards and Technology (NIST), and the Massachusetts Institute of Technology Research and Engineering (MITRE). To facilitate a comprehensive understanding, I will frequently refer to specific cybersecurity terms including Common Vulnerabilities and Exposures (CVE), Common Weakness Enumeration (CWE), and Common Vulnerability Scoring System (CVSS). These terms are essential for my analysis and are defined as follows:

* CVEs are managed by NIST and provide a public method of sharing information about specific cybersecurity vulnerabilities.
* The CWEs database is managed by MITRE and is a category system for types of hardware and software weaknesses and vulnerabilities.
* CVSS is a way to evaluate and rank reported vulnerabilities in a standardized and repeatable way, as determined by the Forum of Incident Response and Security Teams (FIRST).

The primary aim of my project is to explore how CISA's catalog of known exploited vulnerabilities correlates with the information published in NIST and MITRE databases. This catalog, a crucial collection of vulnerabilities frequently targeted by malicious cyber actors, holds significant importance for federal enterprises. Under operational directives, Federal Civilian Executive Branch agencies (FCEB) are mandated to address vulnerabilities listed in this catalog by prescribed due dates (Cybersecurity & Infrastructure Security Agency). I employed several Python3 scripts to connect these disparate data sources, resulting in an enriched known exploitable catalog. Understanding the relationship between these databases is vital, as it could shed light on patterns and trends in cybersecurity threats, potentially guiding future defense strategies and FCEB cybersecurity policy-making efforts. The following sections seek to answer the following question:

*Are certain software weaknesses, as defined by CWEs, more prevalent within CISA's known exploitable catalog than others? If so, what implications arise about commonly exploited software weaknesses, and are specific venders or products more frequently associated with certain CWEs?*

**Data Visualization:**

The results of my investigation are best visualized using histograms given the explorative nature of my hypothesis. Histograms are great for displaying measurements of data. They answer the essential question of 'how much'. This project, at its core, is an exercise in discovery. By analyzing the distribution of CWEs and venders within my enriched CISA known exploitable catalog I aim to uncover relationships between CWEs, venders, and their respective frequencies. As such, histograms are a natural choice for displaying my data and information. The below histograms confirm my initial hypothesis.



Figure 1: Distribution of CWEs across enriched CISA known exploitable catalog. Generated using Excel (Nov 2023)

Figure 1 shows the frequency of each CWE within the enriched catalog. The x-axis lists CWE IDs, while the y-axis shows the count of each CWE. The chart has a clear descending trend, indicating that some CWEs are much more prevalent throughout the catalog than others. For example, CWE-269 appears to be the most common, with counts significantly higher than the others. This suggests that this software weakness is particularly common throughout the data set.

A graph of a number of people

Description automatically generated with medium confidence

Figure 2: Top ten CWEs in enriched CISA known exploitable catalog. Generated using Excel (Nov 2023)

Figure 2 narrows the scope of the data to just the top ten most occurring CWEs in the data set. This specifically highlights the top ten most frequent CWEs. The top three CWEs in the data set map to “Improper Privilege Management’, ‘Input Validation’, and ‘Memory Buffer Overflow’, respectively. This suggests that the most common software weaknesses within the dataset are related to privilege management and input validation. It is important for organizations to properly secure privileges and properly validate input because these are common initial attack vectors used by attackers.

A graph with text overlay

Description automatically generated

Figure 3: Distribution of vendor products across enriched CISA known exploitable catalog. Generated using Excel (Nov 2023)

Figure 3 shows the number of products associated with each vendor that have known vulnerabilities in the catalog. Microsoft has the highest number of products listed, followed by a steep drop-off for other vendors. This suggests that Microsoft products are heavily represented in the catalog. This could be due to its widespread use and popularity, making it a favored target for malicious actors. Especially considering CISAs known exploitable catalog has legal ramifications for FCEBs, it would make sense to assume that Microsoft is heavily used by FCEBs.

A graph with numbers and a bar

Description automatically generated with medium confidence

Figure 4: Top ten vendors in enriched CISA known exploitable catalog. Generated using Excel (Nov 2023)

Figure 4 narrows the scope to the top ten venders within the data set. This chart reiterates that Microsoft products have the most occurrences of known vulnerabilities, followed by a steep drop-off for other vendors like Apple, Cisco, and Adobe. This reinforces the prominence of Microsoft products within the data set.

A chart of a table

Description automatically generated with medium confidence

Figure 5: Heatmap of CWE occurrences for top vendors. Generated using python script in appendix a (Nov 2023)

Maps are useful for displaying the context of which data exist in. Maps of all types are useful at simplifying complexity and allow readers to better visualize the relationships between data. This is why I created the above heatmap. The color contrasts clearly draw the reader's attention to the outliners in the data and can efficiently point out relationships. Figure 5 cross-references the top vendors with the prevalence of different CWEs in their products. The colors indicate the count of occurrences, with darker colors (deep blue, green) indicating higher counts. Microsoft has a significant concentration of CWE-269, which indicates improper privilege management as a pervasive software weakness in Microsoft products listed in CISAs known exploitable catalog.

**Conclusion:**

In summary, the hypothesis that certain software weaknesses are more prevalent within CISAs known exploitable database than others is supported by these visualizations. Certain CWEs, such as CWE-269, and certain vendors' products, especially Microsoft's, show higher counts of known vulnerabilities. This analysis reflects that while vulnerabilities are widespread, some specific types of weaknesses and products are more frequently exploited than others.

**Take Aways:**

This research substantiates the hypothesis that some software weaknesses, defined by CWEs, are more prevalent in CISA's catalog of known exploitable vulnerabilities. Notably, Microsoft emerges as a frequent vendor with a significant share of issues relating to improper privilege management. Microsoft’s large market share as the predominant player in the operating system space likely explain why it is by far the most reoccurring vender. Given this, how can organizations, particularly FCEBs, use this information to better their security posture? For starters, all organizations should follow cybersecurity best practices outlined by industry leaders such as sufficient auditing, enforcing least privilege, and erecting in depth defense throughout their organizations. Identification and securing of high value assets and processes should be paramount for security teams. Development, distribution, and review of business continuity and disaster recovery plans should occur on an ongoing basis. These are well-known, general security practices. This enriched data set contains vulnerabilities that were novel, and unknown before being widely exploited in the real world. This offers a unique perspective on where possible future zero-days may appear, and how organizations may mitigate and discover their use on their own systems. The most common software weakness, improper privilege management, could be addressed by implementing robust auditing on user accounts. Security teams should consider deployment of User and Entity Behavior Analytics (UEBA) solutions. UEBAs use machine learning, algorithms, and statistical baselines to detect anomalies in behavior on corporate and infrastructure devices. Such solutions could detect malicious activity that attempts to, or successfully does deploy a novel attack against organizations. Clearly, such solutions are not direct fixes for CVEs listed on CISAs known exploitable catalog, but solutions such as IBM Security’s QRadar Suite, Rapid7s XDR, or Splunk’s UEBA offer a means for detecting attacks targeting privilege controls.

**Work Cited**

Cybersecurity & Infrastructure Security Agency. (n.d.). Known exploited vulnerabilities catalog.CISA. Retrieved from <https://www.cisa.gov/known-exploited-vulnerabilities-catalog>

**Appendixes:**

**Appendix A: Scrape and Parse from the web**

final\_gen\_heatmap2.py: <https://github.com/blugo2/AIT-664/blob/main/final_gen_heatmap2.py>