Task C

Bryan Lim Jing Xiang (A0233605M)

1 Dataset

1.1 Overview

This dataset consists of CVEs along with their corresponding metrics across the years up to 2023. The data fields of interest are:

Field	Description	Data Type
cve.published	Name (which denotes month too)	String
cve.metrics.cvssMetricV31.cvssData.baseSeverity	Severity of CVE	String
cve.metrics.cvssMetricV31.cvssData.attackVector	Attack Vector of CVE	String
Year	Year	Integer

Naturally, such a dataset is interesting in view of the recent news over the XZ exploits.

1.2 Data origin

https://www.kaggle.com/datasets/manaielyes/cve-vulnerabilities

1.3 Github Repository

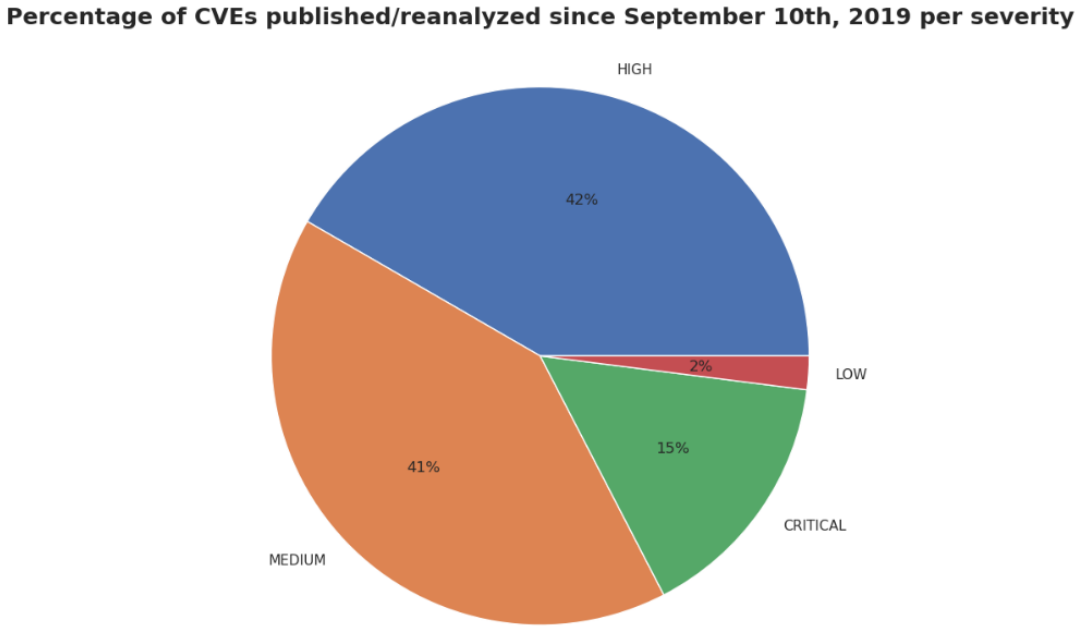
https://github.com/bryanljx/visualisation

2 Purpose of Visualisation

For this dataset, the query of interest is: "What are current trends/patterns amongst CVEs nowadays? How severe are they and what kind of vulnerabilities are out there?"

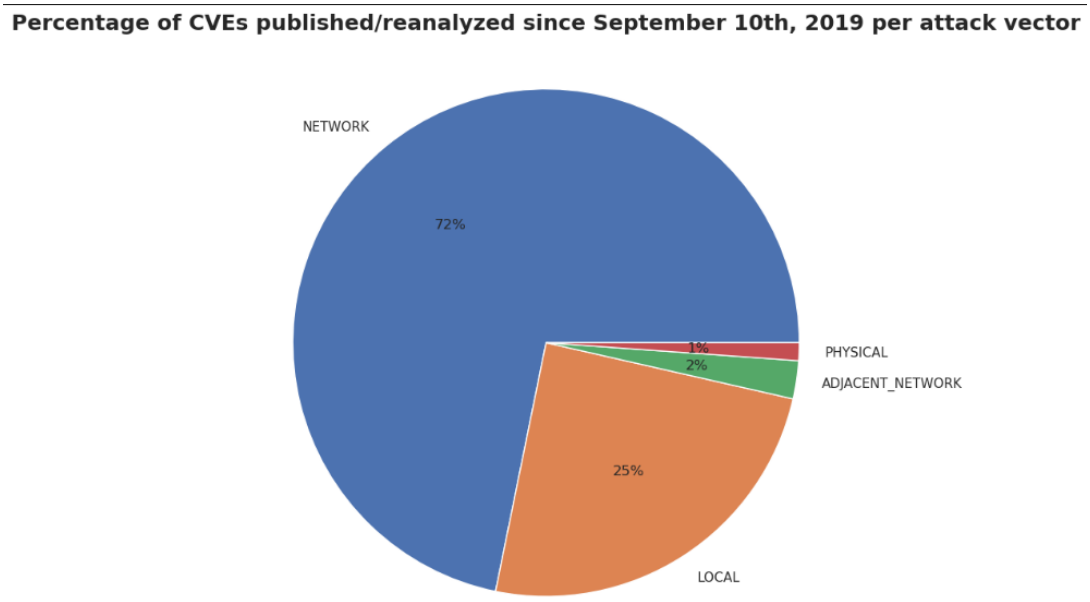
3 Visualisation

3.1 CVEs per severity category since September 10th, 2019



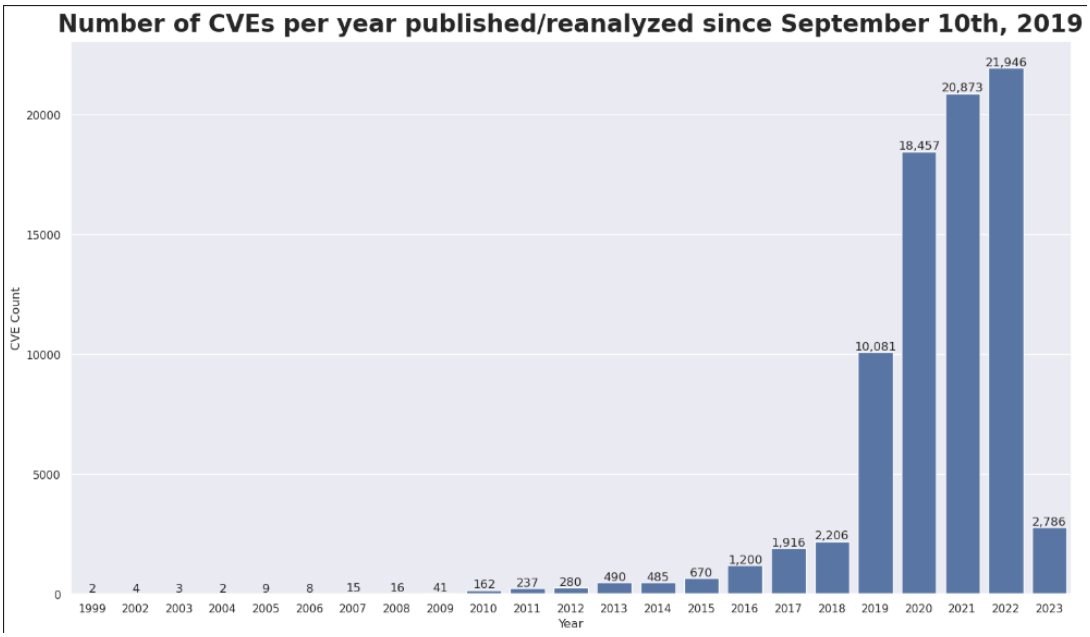
- Note: Only data points which have values for the new CVSS v3.1 metric are included, this means only CVEs published after this data or reanalysed after this data are included, which makes for a more recent and current view of CVEs that are of concern today.
- A pie chart was chosen here to show the distribution of CVEs per severity.
- Visual encoding here includes:
  - Length - Denoting the percentage of CVEs for each severity
  - Color - Denoting the severity
- Insights
  - Most CVEs are of medium or high severity, with only a select few being critical.

3.2 CVEs per attack vector since September 10th, 2019



- A pie chart was chosen here to show the distribution of CVEs per attack vector.
- Visual encoding here includes:
  - Length - Denoting the percentage of CVEs for each attack vector
  - Color - Denoting the attack vector
- Insights
  - Network is the most vulnerable attack vector

3.3 Distribution of current CVEs across the years



- A bar chart was chosen to show the number and spread of CVEs still relevant (only those reanalysed or published after September 10th, 2019) across the years
- Visual encoding here includes:
  - Length: Denoting the number of CVEs published that year still relevant now
- Insights
  - Relatively huge number of CVEs from 2019 - 2022. While this coincides with a greater amount of CVEs reported and published, this may suggest that CVEs are not tackled and resolved that fast, that they are still relevant and a potential attack vector for a few years.