Why is Cybercrime Under-reported and do Cyber Security Incidents in UK Organisations Result in ICO Enforcement?

Business Analytics and Management

A Dissertation submitted to the Norwich Business School of the University of East Anglia in partial fulfilment of the requirements for the degree of Master of Science.

SUPERVISOR(S), MARKERS/CHECKER AND ORGANISER

The undersigned hereby certify that the markers have independently marked the dissertation entitled "Access Granted: On the Security of Near-Field Enabled Keycards" by Stephen Jones, and the external examiner has checked the marking, in accordance with the marking criteria and the requirements for the degree of Master of Science.

Supervisor: ______________________________

Second Marker: ______________________________

External Examiner: ______________________________

Moderator: ______________________________

Dr. Stephen J. Jones ______________________________

DISSERTATION INFORMATION AND STATEMENT

Dissertation Submission Date: 31 August 2023

Student Number: 100383119

Title: Why is Cybercrime Under-reported and do Cyber Security Incidents in UK Organisations Result in ICO Enforcement?

School: Norwich Business School

Course: Business Analytics and Management

Degree: MSc

Year: 2023

Organiser: Dr. Stephen J. Jones

STATEMENT:

Unless otherwise noted or referenced in the text, the work described in this dissertation is my own work to the best of my knowledge and belief. It has not been submitted, either in whole or in part for any degree at this or any other academic or professional institution.

Abstract

In the world dominated by digital transformation, understanding cybercrime's intricacies has never been more crucial. This dissertation looks deep into the cyber realm, addressing critical problems like cybercrime under-reporting and investigating the efficacy of cyber security incident responses, notably those provided by the Information Commissioner's Office (ICO) in the United Kingdom. The investigation began with identifying inconsistencies in cybercrime statistics, particularly between Action Fraud (AF) and ICO. These variations not only raise eyebrows, but also call into doubt the data's legitimacy and dependability. The study uncovers patterns and trends by combining historical data spanning several years with predictive modelling. These findings are critical in that they provide light on developing cyber threats and the potential limitations of predictive algorithms in the face of constantly shifting cyber attackers. The literature review, which offers a thorough picture of the current body of knowledge, is a critical element of the dissertation. The literature provides an in-depth analysis of cybercrime, from its definition to its classification to the motives behind these digital offences. The analysis highlights phishing as a common mode of operation, with attackers using sophisticated strategies to trick unsuspecting victims. Similarly, unauthorised access appears as a major problem, with several possible consequences ranging from financial consequences to trust decline. However, the analysis highlights the apparent under-reporting of cybercrime. The underlying causes are numerous, including a lack of understanding, mistrust in the reporting system, potential consequences, and, in some cases, pure misunderstanding of being a cyber victim. Addressing under-reporting is critical not just for accurate data representation but also for developing effective responses. The methodology section describes the methodical approach used, with an emphasis on data gathering, analysis, and interpretation. Predictive modelling, particularly linear regression, becomes the technique of choice for forecasting cyber dangers in the future. The study does, however, emphasise the inherent limits of such models, especially in a changing digital context. The results and discussion sections represent the core of the dissertation, presenting and analysing the findings. The gaps between AF and ICO data are obvious, producing more concerns than solutions. While trends and patterns emerge, the data also exposes the dangers of depending simply on previous data to forecast the future. The digital domain is constantly changing, driven by technology advancements, global events, and legislative shifts, making cybercrime prediction a difficult undertaking. The research concludes in recommendations, highlighting data harmonisation, focusing efforts on prevalent risks such as malware, improving unauthorised access reporting, and refining predictive modelling. Collaboration, both intra- and inter-agency, is being praised as the way ahead to ensure an integrated front against cyber attackers. The study also promotes public awareness, arguing for an informed digital citizen capable of recognising and reporting threats. Furthermore, the role of the IT industry is emphasised, emphasising the importance of a mutually beneficial relationship between cyber agencies and the tech world. Finally, the dissertation provides a complete assessment of the cyber world, addressing the issue of under-reporting and analysing the ICO's enforcement efforts. The results and recommendations are practical tools for policymakers, cyber organisations, and the general public, ensuring a safer digital future for all.

Acknowledgements

This dissertation marks the culmination of my studies for the degree of Master of Science. I thank the Norwich Business School for what has been both a challenging and enjoyable experience. I also wish to acknowledge and thank the following outstanding individuals, who have provided great support and encouragement throughout this project:

- Dr. Stephen J. Jones
- Ms. Diksha Barad
- Mr. Muhammad Faheem
- Mr. Rahul Gupta

Dedicated to my Mamma and Dadda

Contents

List of figures

1.Introduction

Cybercrime is a threat brought on by a series of technical improvements, and it presents considerable difficulties for UK organisations, law enforcement agencies, and regulatory bodies. The focus of this dissertation is the relationship between reporting on cybercrime, regulatory enforcement, and the subsequent cybersecurity impact.

1.1 Dissertation Structure

The organised structure of this dissertation allows for a thorough study of its main idea. Beginning with a summary of the widespread problem of underreported cybercrime before moving on to a thorough literature study that examines crime reporting trends and regulatory changes. The next step is data analysis, which seeks answers to crucial research issues and ends with practical suggestions for UK organisations. Each chapter is painstakingly crafted to lead readers through a clearly written story while providing insight, clarity, and useful information.

1.2 Background and Motivation

Substantial academic and professional attention has been paid to the topic of underreported cybercrime in the UK (McGuire & Dowling, 2013; Wolff, 2018). The consequences of this underreporting are severe and hinder attempts by law enforcement agencies to prevent cybercrime before it happens. Without thorough reporting, police forces are in the dark and unable to plan effective defences. However, the issues don't stop here. The Data Protection Act of 2018—which was in line with the game-changing EU GDPR—highlights the requirement that UK organisations swiftly report the

ICO of cybersecurity breaches that expose personal data. This act marked a paradigm change, increasing the scrutiny surrounding the handling of personal data and the severity of the penalties for its abuse. These changes, together with the requirement to safeguard organisational assets and reputation, fuel the motivation for this research.

1.3 Objectives

Business Objective

From a business point of view, the goal goes beyond simple compliance. Businesses must understand the importance of underreporting cybercrimes given the consequences of data breaches, including the financial penalties and reputational harm. Organisations need to be extra certain about their reporting procedures and cybersecurity precautions because of the

Data Protection Act (2018), which further tightens the noose. To give organisations useful information that might direct their cybersecurity strategy and reporting procedures in the face of potential threats, this research seeks to clarify the motivations and boundaries of reporting.

Technical Objective

On the technological front, it is hoped to use advanced analytical methods and investigate deeply into the datasets that are currently available, particularly from Action Fraud and the ICO. The goal is to determine correlations between Action Fraud and ICO data on cybercrime, evaluate the frequency and effectiveness of enforcement actions taken by ICO after data breaches, and use predictive analytics to identify trends in cybercrime. Achieving these goals will provide a detailed picture of the cybercrime ecosystem today and set the stage for future technology defences and reporting guidelines.

2. Literature Review

It takes a thorough analysis of the body of available literature to comprehend the complex nature of cybercrime and its reporting procedures inside the UK. This discussion will focus on Action Fraud's reporting mechanisms, the function and enforcement powers of the Information Commissioner's Office (ICO), and a thorough examination of the numerous types of cybercrime that are common in the digital era.

2.1 Action Fraud

The role of Action Fraud and the complexities that go along with it are critical in the context of this study because it has made a name for itself as the main reporting point for fraud and cybercrime in the UK. The importance of this body has grown recently, especially considering the rise in sophistication and frequency of cybercrimes (Giro Correia, 2022). Studies have found varying degrees of success in getting victims, particularly organisations, to come forward through Action Fraud's reporting procedures and outreach activities (Beshi & Kaur, 2019). This organisation's data collection offers insights into the trends, tactics, and underreporting aspect of cybercrimes. Contends that there is still potential for improvement, notably in terms of public awareness, even if Action Fraud played a key contribution in centralising crime reporting (Smith, 2000).

2.2 Information Commissioner’s Office (ICO)

The General Data Protection Regulation (GDPR) of the EU and the Data Protection Act of 2018 are two significant pieces of legislation that organisations must abide by, and the ICO is essential in ensuring that they do. The adoption of these legal tools signified a more difficult attitude to data privacy in the UK (Bhaimia, 2018). Much discussion has centred on the potential for the ICO to hold businesses accountable for violating data protection regulations, particularly through the enforcement of substantial fines. Some claim that although many see these enforcement agencies as obstacles, they may unintentionally discourage organisations from reporting violations (Robinson & Graux, 2009). An examination of the ICO's position following the implementation of the GDPR reveals both its advantages and disadvantages in preserving data privacy at a time when cyberattacks are common (Linden et al., 2019).

2. 3 Computer Misuse Act 1990

A ground-breaking piece of legislation, the Computer Misuse Act 1990, was established in the UK to address the growing concerns over computer security, unauthorised access, and other computer-related offences (Parliament, 1990). Prior to the passage of the act, the UK legal system struggled to prosecute computer crimes since there were no explicit statutes addressing these offences.

Both in the business world and among the public, computer use increased dramatically in the 1980s. With the emergence of the digital age came a new breed of crimes, from breaking into secure networks to distributing dangerous malware. There was a clear need for a comprehensive legal framework to address these issues, which led to the establishment of the Computer Misuse Act **(Great Britain. Her Majesty's Stationery Office, 1990).**

The act is structured around three primary offences:

1. Unauthorised Access to Computer Material: The act's specific provision that makes accessing computers without the necessary authorization illegal. Given the widespread adoption of passwords and cybersecurity, such an offence may appear simple by modern standards. In the early days of computing, however, even minor instances of unauthorised access might result in serious data breaches or substantial financial losses. This clause set the stage for criminal prosecution of those who, for example, used someone else's password to obtain unauthorised access to a (Parliament, 1990).

2. Unauthorised Access with Intent to Commit or Facilitate Commission of Further Offenses: This section relies on the first provision's foundation by introducing a higher level of obligation. It focuses not just on unauthorised access, but also on the motivation behind such activities. If a person gains access to a system with the goal to conduct another crime, such as fraud or data theft, they may be punished under this harsher provision. (Parliament, 1990).

3. Unauthorised Acts with Intent to Impair, or with Recklessness as to Impairing, Operation of a Computer: This section focuses on the most serious forms of cybercrime. It includes actions that are intended to cause damage or harm computer systems or networks, such as the introduction of viruses or other malware. The addition of' recklessness' broadens the scope, guaranteeing that even individuals who act without open intent but with disregard for potential harm can be held accountable. (Parliament, 1990).

The enactments of the Computer Misuse Act were met with varied reactions. While many praised it for filling a void in the legal system, others criticised it for being overly wide or unclear in its definitions, which could lead to misinterpretation or misuse (McGuire, 2007).

The law has been modified over time to address the continuously changing reality of cybercrime. The Police and Justice Act of 2006 and the Serious Crime Act of 2015 are two significant instances of legislation that introduced changes to better address the modern difficulties provided by cybercriminals (Clough, 2015).

Finally, the Computer Misuse Act of 1990 was a watershed point in the United Kingdom's approach to computer-related offences, establishing a framework that has subsequently been expanded and enhanced.to tackle the multifaceted challenges of cybercrime.

Extortion:

Extortion, a timeless crime, involves coercing someone into providing money, property, or services under duress. While historically associated with organized crime and blackmail schemes, extortion has evolved dramatically in the digital age, leveraging technology to target individuals, corporations, and even governmental entities (Maimon et al., 2013).

Digital extortion can be seen prominently in the form of ransomware attacks. Here, attackers compromise a victim's data or system, demanding payment, often in cryptocurrencies, to restore access (Ali, 2017). However, ransomware is merely one mechanism. Cyber extortion can also involve threats to disclose sensitive information, or to launch a crippling distributed denial-of-service (DDoS) attack against an entity's digital infrastructure (Bendovschi, 2015).

The motivations behind digital extortion are diverse. While financial gain remains primary, political motivations, revenge, or merely the thrill of the act can also drive perpetrators (Ali, 2017). The advent of cryptocurrencies like Bitcoin has further incentivized cyber extortion due to the perceived anonymity these platforms offer (Iriberri & Navarrete, 2010).

Protection against digital extortion necessitates a multi-faceted approach. Regular data backups, cybersecurity training, strong security protocols, and timely software updates are foundational. Legally, reporting extortion attempts, even if they seem baseless This not only aids in potential prosecution but helps security agencies and experts to identify trends, tools, and tactics used in the ever-evolving landscape of cyber threats (Buckley, 2021).

Malware/ Ransomware:

Malware, a term derived from 'malicious software', encompasses a broad category of software intentionally designed to harm, exploit, or otherwise compromise the confidentiality, integrity, or availability of computer systems (Houser, 2015). One of the most malicious forms of malware that has seen a significant surge in the past decade is ransomware.

Ransomware is a subtype of malware that restricts access to the infected system, and the data therein, demanding a ransom be paid to the perpetrator to restore access (Al-rimy et al., 2018). It functions by encrypting victim's files and demanding payment, typically in cryptocurrency, for decryption keys. The notorious 'WannaCry' outbreak of 2017 is a testament to the global threat that ransomware can pose, affecting over 230,000 computers across 150 countries and causing disruptions in critical sectors like healthcare and transportation (Chen & Bridges, 2017).

Two primary types of ransomwares exist: crypto ransomware, which encrypts valuable files, and locker ransomware, which locks the user out of their device (Sgandurra et al., 2016). The evolution of ransomware has been driven by several factors, including the rise of cryptocurrencies which provide anonymous transaction capabilities, making it difficult for authorities to trace cybercriminals (Kharraz et al., 2015).

The consequences of a ransomware attack are multifaceted. Beyond the immediate financial implications of paying a ransom, victims may suffer data loss, business disruption, reputational harm, and potential legal and compliance ramifications (FBI, 2016). Furthermore, there's no guarantee that payment will result in data recovery.

Preventing ransomware attacks primarily revolves around robust cybersecurity hygiene. Regular backups, updated software, user education, and advanced threat detection tools form the crux of a proactive defence strategy (Radanliev et al., 2019). The human element is particularly vital: with many ransomware attacks initiated through phishing tactics, training individuals to recognize and report suspicious activity is crucial (Maniath et al., 2019).

Phishing:

Phishing remains one of the most prevalent forms of cyber threats, characterized by deceptive attempts to acquire sensitive information such as usernames, passwords, and credit card details by masquerading as a trustworthy entity (Jagatic et al., 2007). By leveraging the human element, cybercriminals exploit psychological mechanisms, like trust and fear, to deceive victims into revealing confidential information or executing unwanted actions.

The term "phishing" finds its origin in the word "fishing", symbolizing the act of casting a baited hook to see who bites (Emigh, 2006). The evolution of this cyber-attack method over time has been notable. Initially, phishing attacks were simple, relying on broad spamming methods. However, they have become increasingly sophisticated, utilizing personalized information to make attacks more convincing—a method termed "spear phishing" (Zhang et al., 2018).

The consequences of phishing can be catastrophic for both individuals and organisations. For individuals, this could translate into unauthorised transactions, identity theft, and potential financial loss (Workman, 2008). Businesses, on the other hand, can suffer reputational damage, financial losses, and legal ramifications due to data breaches arising from successful phishing attacks (Lastdrager, 2014).

The strategies used by phishers often mirror real-world con artistry. These attackers use cues to appear legitimate, such as mimicking authentic logos, using official-sounding language, or leveraging urgency by inducing panic or immediate action (Kumaraguru 2007).

While technology-based solutions, such as spam filters and email authentication, can help reduce the risk, training and educating end-users is pivotal. Human error remains the most significant vulnerability in phishing attacks, highlighting the importance of raising awareness and building a culture of suspicion regarding unsolicited communications (Alotaibi et al., 2016).

Unauthorised access:

Unauthorised access, a significant form of cyber violation, occurs when individuals gain illegal entry into systems, networks, or databases without permission (Kianpour, 2021). This breach is one of the most conventional cybercrimes and poses a substantial threat to individuals and organisations alike.

The digital landscape has always been fraught with vulnerabilities that skilled hackers can exploit. Digital systems, regardless of their sophistication, inevitably have weak points that can become targets for malicious actors. In (Rai, 2010) argue that the very nature of digital platforms, with complex interdependencies, creates a breeding ground for potential access points for hackers. The motivation behind such breaches can vary from the thrill of the act itself, known as "hacktivism," to more malicious intents like data theft, espionage, or system disruption (Furnell & Warren, 1999).

The ramifications of unauthorised access are diverse and, in many cases, devastating. For organisations, such a breach can led to significant financial losses, especially if sensitive data such as customer details or proprietary information is accessed (Anderson & Moore, 2006). Beyond financial implications, there's also a loss of trust from stakeholders and customers, a factor which might take years to rebuild, if at all. Ethical concerns around privacy and the autonomy of personal information are significant issues in the digital age (Tavani, 2007).

The increasing integration of IoT (Internet of Things) devices has added another layer of complexity to the issue. With more devices connected to the internet, the potential points of unauthorised access have multiplied (Harit et al., 2017). These devices, often without robust built-in security measures, can become easy gateways for hackers seeking access to larger systems or networks.

To counteract unauthorised access, a multi-layered approach to cybersecurity is crucial. Employing intrusion detection systems, regular audits, encryption, and comprehensive cybersecurity training for employees are just some measures that organisations can implement. However, as technologies evolve, so do hacking methodologies. As a result, staying ahead requires continuous vigilance, updated knowledge, and proactivity in system protection.

3. Methodology

3.1 Overview

The amount of data quality that is used in any research has a significant impact on its reliability. Due to the complexity of the cybercrime landscape, the datasets in question, which provide quarterly results and contain a wide range of attributes, each of which indicates a specific sort of cyber-crime incidence, are essential. Accurate data cleaning and preparation process was needed because the main objective was to focus on cybercrimes that hurt businesses. This section explains the organized method used to clean up and ready the data for additional analysis.

Research Questions:

1. What association or correlation is there between the Action Fraud and ICO Cybercrime Data?

2. Is there sufficient evidence that successful ICO enforcement action and prosecution can be associated or correlated with cyber security breaches?

3. To what extent can future cybercrime and fraud be predicted using the available historic data?

3.2 Data Preparation

1. **Data Assessment:**

To have a full understanding of the variables and the data's structure, a preliminary review of the datasets was first conducted. It was possible to spot any apparent irregularities, discrepancies, or missing values as a result to this procedure.

The dataset's key properties were found to be hacking (personal and email), ransomware, extortion, unauthorised access, denial of service, and other cybercrime occurrences.

2. **Data Segregation for Business-focused Analysis:**

Due to the study's focus on businesses, entries that were unrelated to cybercrimes against businesses were removed. By streamlining the dataset, the next stages of analysis may be completely linked with the study's goals. When cybercrimes had uncertain classifications, they were reclassified using the dataset's thorough descriptions to ensure appropriate representation in categories pertaining to business.

3. **Data Consistency Check:**

Because the data was given quarterly, it was critical that the date formats remained consistent throughout. This allowed for straightforward temporal analysis. Instead of quarterly statistics, yearly data were used instead. Given the range of cybercrimes mentioned, it was critical to ensure that attribute labels were consistent. Any contradictions, such as the use of several terminology for the same type of offence, were corrected.

3.3 Data Staging

1. **Data Normalization:**

Data normalisation procedures were used to guarantee that all attributes were on a comparable scale, which was especially crucial for subsequent statistical analysis. This guaranteed that one attribute did not have an excessive impact on the outcomes due to its size.

2. **Data Integration:**

Data has been taken from databases containing information on various types of crimes committed around the United Kingdom. It was then put into a framework that provided consistency in terms of quality and dates. When different sources employed different terminologies, an attribute mapping approach was created to allow for the seamless merger of datasets. Random samples were extracted and examined from the cleaned dataset to check correctness, consistency, and relevance.

3.4 Descriptive Statistics

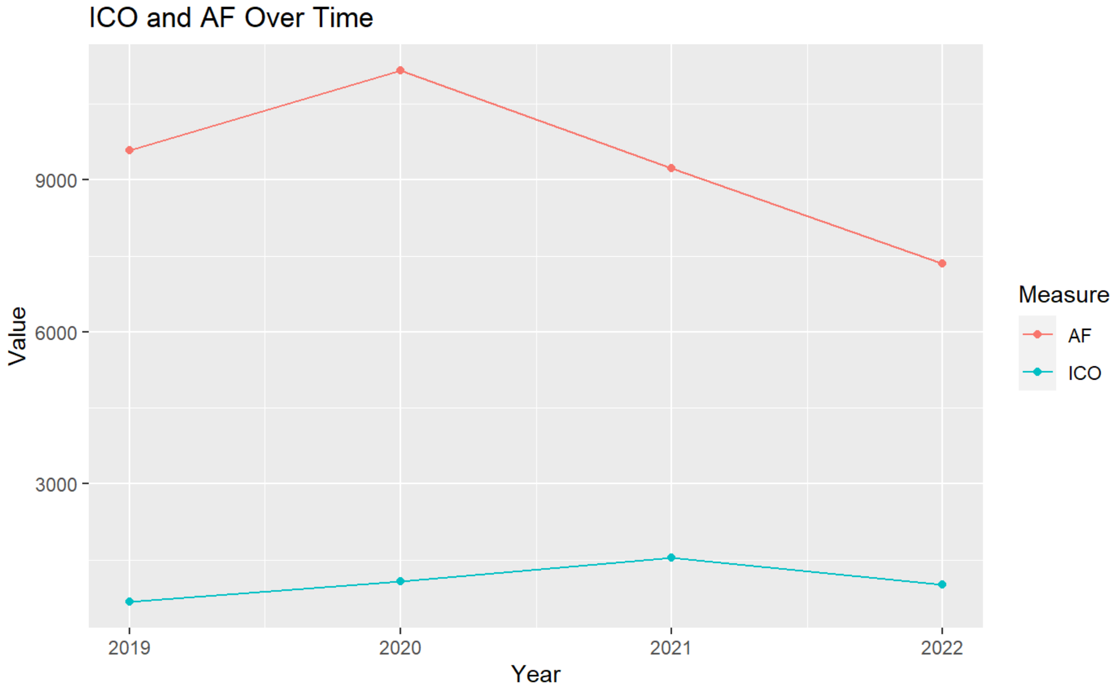
Descriptive statistics were generated for the cleaned dataset. This provided an overview of the data distribution and aided in identifying any remaining outliers or anomalies.

```
#Define the data
year <- c(2019, 2020, 2021, 2022)
ICO <- c(678, 1071, 1542, 1019)
AF <- c(9589, 11158, 9231, 7344)
# Create a data frame
data <- data.frame(year, ICO, AF)
# Convert from wide format to long format
data_long <- data %>% pivot_longer(-year, names_to = "Variable", values_to = "Value")

# Summary of the data
summary(data)
```

```
    year          ICO             AF
 Min.   :2019   Min.   : 678.0   Min.   : 7344
 1st Qu.:2020   1st Qu.: 933.8   1st Qu.: 8759
 Median :2020   Median :1045.0   Median : 9410
 Mean   :2020   Mean   :1077.5   Mean   : 9330
 3rd Qu.:2021   3rd Qu.:1188.8   3rd Qu.: 9981
 Max.   :2022   Max.   :1542.0   Max.   :11158
```

```
# Plot the data
ggplot(data_long, aes(x = year, y = Value, color = Variable)) +
geom_line() +
geom_point() +
labs(x = "Year", y = "Number of crimes",
title = "ICO and AF Over Time") +
```



```
scale_color_discrete(name = "Measure") # Changing the legend title
```

Fig 1: Descriptive statistics

Any data analysis starts with descriptive statistics as its pillar. Without having to study the data in its full, these statistics offer a brief overview of the key elements. The ability to see things from above is extremely helpful when working with vast datasets, such as those that reveal trends over time.

The two variables ICO and AF, tracked over a four-year period from 2019 to 2022, are the focus of the analysis. These variables descriptive statistics are important because they put light on how they behave and are distributed across time.

1. **The Type of Data**

Understanding the data's nature is crucial before looking into the statistics. Given that ICO and AF deal with numerical counts, the data is in the interval or ratio scale. Most statistical measures, such as spread and central tendency measures like the mean and standard deviation, work well with this type of data.

2. Summary Function

The summary function in R provides a quick and comprehensive snapshot of the data. For both ICO and AF, it calculates:

Minimum (Min.): This indicates the smallest value for each variable in the dataset. For instance, the minimum value for ICO over the course of four years is 678, whereas for AF it is 7344 in the snippet presented. When considering a range or identifying anomalies, the minimum value might provide information about the lowest recorded cases or counts for certain variables.

1st quarter (1st Qu.): The 25th percentile of the data is shown by this value. Alternatively stated, 25% of the data points fall below this figure. It is a metric that can be used to pinpoint the data's lower range.

Minimum (Min.): This designates the smallest value for each variable in the dataset. For instance, the minimum value for ICO over the course of four years is 678, whereas for AF it is 7344 in the snippet presented. When considering a range or identifying anomalies, the minimum value might provide information about the lowest recorded incidences or counts for certain variables.

1st Quartile (1st Qu.): The 25th percentile of the data is shown by this value. Alternatively stated, 25% of the data points fall below this figure. It is a metric that can be used to pinpoint the data's lower range.

Mean (Mean): The dataset's mean is the average of all the values. Although it offers a centre value, outliers can have an impact on it. The mean provides an average count of incidents for ICO and AF over the years in the analysis that is provided.

3rd Quartile (3rd Qu.): This value denotes the 75th percentile and implies that 75% of the data points fall below it. It facilitates comprehension of the data distribution's top range.

Maximum (Max.): The highest value in the dataset is the maximum value. It stands for the year with the highest occurrence or count totals for ICO and AF in the context of the analysis.

The 'ICO' variable has a minimum value of 678, signifying the lowest number of incidents recorded in a year, and a maximum value of 1542 incidents. The mean (or average) of the 'ICO' values is 1077.5, indicating that there were around 1078 events each year across the investigated years. The first and third quartiles (933.8 and 1188.8) provide a 50% range or interquartile range, showing that the middle 50% of the data lies between these numbers.

Within these years, the data provides a minimum of 7344 events and a maximum of 11158 incidents for the 'AF' variable. An average of 9330 events occurs each year. The interquartile range is between 8759 (1st quartile) and 9981 (3rd quartile) events, reflecting the spread of the 'AF' data's middle 50%.

Descriptive statistics, provides context for the data. The difference between the mean and median, for example, indicates the irregularity of the data. A big difference between the first and third quartiles implies that the data is more variable, whereas a minor difference may imply that the data is more consistent.

Understanding these ICO and AF numbers in the context of the analysis provides insight into several different topics. Over time, have incidences been consistently high or low? Years with sharp increases or decreases have there been. The answers to these queries can direct future investigation, such as searching for potential outside influences on such changes.

3.5 Correlation

3.5.1 T- test

When two samples have different variances and sample sizes, Welch's t-test, also known as the unequal variances t-test, is used to determine whether two population means are different. It is a variation of the independent two-sample t-test and is more accurate when the variances and/or sample sizes of the two samples are different.

The main difference between the Welch's t-test and the traditional t-test is its degree of freedom computation, which adjusts based on sample variances and sizes to produce findings that are more accurate when inequalities is present.

```
com_mean <- read_xlsx("data/CORR.xlsx")
clean_names(com_mean)

# A tibble: 8 × 4
   year type  malware   uaa
  <dbl> <chr>   <dbl> <dbl>
1  2019 AF       6745   311
2  2020 AF       7618   339
3  2021 AF       6412   286
4  2022 AF        433   199
5  2019 ICO        75   473
6  2020 ICO       101   597
7  2021 ICO       134   619
8  2022 ICO        70   175

t.test(com_mean$Malware~com_mean$Type, data = com_mean)


    Welch Two Sample t-test


data:  com_mean$Malware by com_mean$Type
t = 3.1695, df = 3.0005, p-value = 0.05049
alternative hypothesis: true difference in means between group AF and group ICO is not equal to 0
95 percent confidence interval:
   -20.84869 10434.84869
sample estimates:
 mean in group AF mean in group ICO
             5302                95

t.test(com_mean$UAA~com_mean$Type, data = com_mean)


    Welch Two Sample t-test


data:  com_mean$UAA by com_mean$Type
```

```
t = -1.7102, df = 3.5219, p-value = 0.172
alternative hypothesis: true difference in means between group AF and group ICO is not equal to 0
95 percent confidence interval:
 -494.6665  130.1665
sample estimates:
 mean in group AF mean in group ICO
          283.75            466.00

# Load necessary packages
library(ggplot2)

# Plot for Malware
p1 <- ggplot(com_mean, aes(x=Type, y=Malware)) +
  geom_boxplot() +
  stat_summary(fun=mean, geom="point", shape=20, size=3, color="red") +
  labs(title="Boxplot of Malware by organisation type",
       y="Malware",
       x="Organisation_type")

# Plot for UAA (Unauthorised Access Attempts)
p2 <- ggplot(com_mean, aes(x=Type, y=UAA)) +
  geom_boxplot() +
  stat_summary(fun=mean, geom="point", shape=20, size=3, color="red") +
  labs(title="Boxplot of Unauthorised access by organisation type",
       y="Unauthorised_access",
       x="Organisation_type")

# Print plots
print(p1)
```

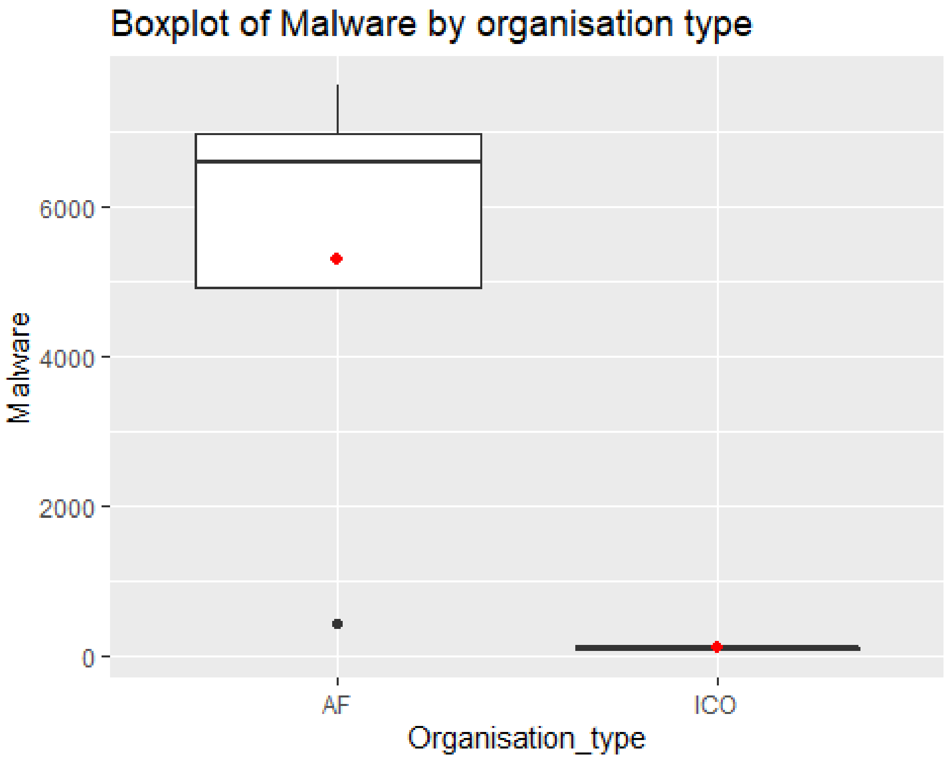


Fig 2: Box plot for malware

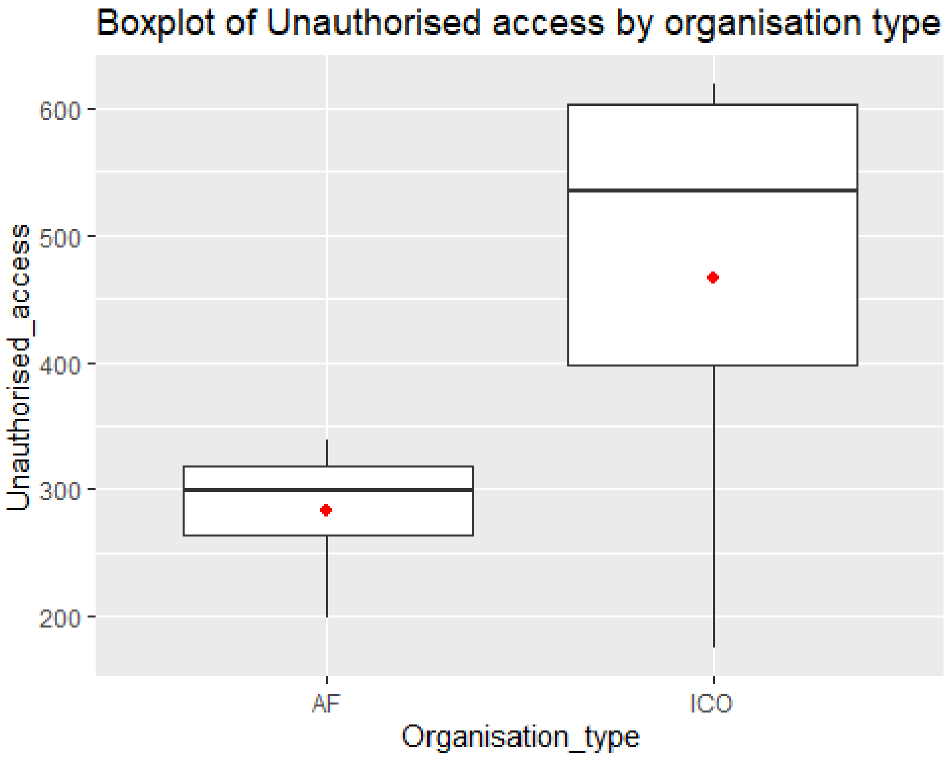


Fig 3: Box plot for unathorised access

T-test: From both the dataset, I have picked Malware and Unauthorised access attributes as they both were common in both the datasets.

Test for Malware: The test compares the mean values of the variable ‘Malware’ between two groups: AF and ICO.

Test Statistic (t): The t-value is 3.1695. This value measures the difference between the group means in terms of standard errors. A larger t-value indicates a larger difference between the groups.

Degrees of Freedom is 3.0005. This is a measure of the sample size used in the test, adjusted for the variances of the two groups. Since this is a Welch t-test, which does not assume equal variances.

P-value: The p-value is 0.05049. Typically, a p-value less than 0.05 is considered evidence that there is a statistically significant difference between the groups. This p-value is right on the edge, meaning there is evidence for a significant difference.

Confidence Interval: The 95% confidence interval for the difference in means ranges from -20.84869 to 10434.84869. Since this interval contains 0, it indicates that the true difference between the group means could be zero, although the range is quite large.

Sample Estimates: The mean value of ‘Malware’ for group AF is 5302, while for group ICO it is 95. This is a substantial difference in raw terms.

Test for Unauthorised Access: The test compares the mean values of the variable Unauthorised access between the groups AF and ICO.

Test Statistic (t): The t-value is -1.7102. The negative sign indicates that the mean of group AF is lower than that of group ICO.

Degrees of Freedom is 3.5219.

P-value: The p-value is 0.172. This is greater than 0.05, indicating that there is not enough statistical evidence to say that the means of the two groups are different.

Confidence Interval: The 95% confidence interval for the difference in means ranges from -494.6665 to 130.1665. Since this interval contains 0, it supports the idea that there might not be a significant difference between the two group means.

Sample Estimates: The mean value of Unauthorised access for group AF is 283.75, while for group ICO it is 466.00.

Summary: For the ‘Malware’ variable, there seems to be a substantial difference in raw terms between the two groups, with the AF group having a much higher mean. However, the statistical test provides only borderline evidence for this difference.

For the ‘Unauthorised access’ variable, while there is a difference between the means of the two groups in raw terms, the statistical test does not provide evidence that this difference is significant.

3.5.2 Correlation Matrix

```
# Create the datasets
ICO <- data.frame(
  Year = c(2019, 2020, 2021, 2022),
  Malware = c(75, 101, 134, 70),
  Other_cyber_incidents = c(53, 213, 260, 232),
  Ransomware = c(77, 160, 529, 542),
  Unauthorised_access = c(473, 597, 619, 175),
  Sum = c(678, 1071, 1542, 1019)
)

AF <- data.frame(
  Year = c(2019, 2020, 2021, 2022),
  Malware = c(6745, 7618, 6412, 4333),
  Unauthorised_access = c(311, 339, 286, 199),
  Extorsion = c(2533, 3201, 2533, 2812),
  Sum = c(9589, 11158, 9231, 7344)
)

# Merging datasets by 'Year'
merged_data <- merge(ICO, AF, by = "Year")

# Drop 'Year' column for correlation calculation
cor_data <- merged_data[,-1]

# Compute the correlation matrix
cor_matrix <- cor(cor_data)
```

```
corrplot(cor_matrix, method = "circle", type = "upper", order = "hclust",
        tl.col = "black", tl.srt = 45)
```

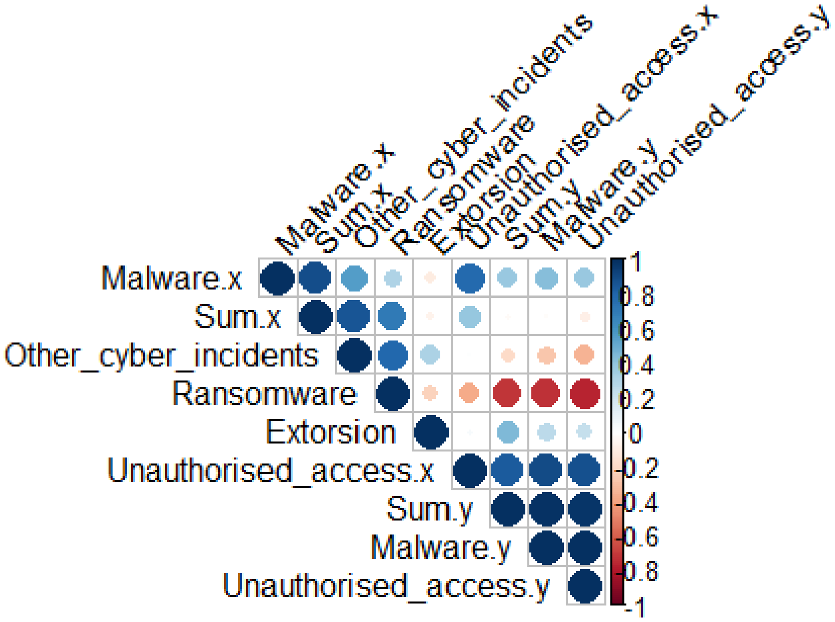
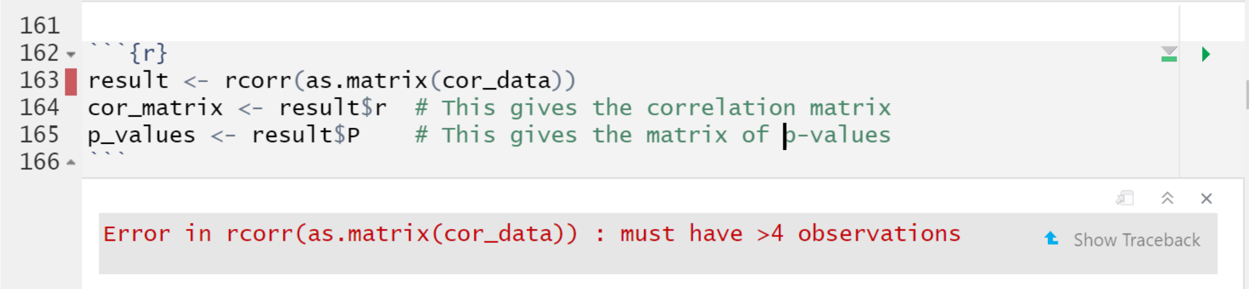


Fig 4: correlation plot

From the above correlation matrix, we cannot find the P value as the data should have more than 4 years of data and we have only 4 years of data from the ICO. (Attached the error below, with the code).



~~~
```{r}
result <- rcorr(as.matrix(cor_data))
cor_matrix <- result$r   # This gives the correlation matrix
p_values <- result$P     # This gives the matrix of p-values
```
~~~

```
Error in rcorr(as.matrix(cor_data)) : must have >4 observations
```

ICO Dataset:

Year: Spans from 2019 to 2022.

Incident Categories: Malware, Other Cyber Incidents, Ransomware, Unauthorised Access.

Sum: Represents a cumulative count of all incidents for the year.

AF Dataset:

Year: Covers the same years, 2019 to 2022.

Incident Categories: Malware, Unauthorised Access, Extortion.

Sum: A cumulative tally of incidents for each year.

The datasets were subsequently merged based on the 'Year', resulting in a comprehensive dataset called merged data.

Correlation Analysis

The primary objective of this analysis is understanding the relationships between types of cyber incidents. By computing the correlation matrix using the cor function, we derive a numerical representation of these relationships. The range of correlation values lies between -1 and 1, with:

-1 indicating a perfect negative correlation.

0 suggesting no correlation.

1 denoting a perfect positive correlation.

the inability to compute the P-value due to the dataset spanning only four years. P-values are crucial in determining the statistical significance of observed correlations. Without them, while we can comment on relationships, we cannot confidently assert their statistical significance.

3.6 Regression Analysis for prediction

3.6.1 Regression model for Action Fraud:

```
# Linear Regression analysis for Action Fraud
# Using the dataset you provided:
AF <- data.frame(
Year = c(2015,2016,2017,2018,2019, 2020, 2021, 2022),
Malware = c(3568,7259,6691,4177,6745, 7618, 6412, 4333),
Unauthorised_access = c(521,691,689,708,311, 339, 286, 199),
Extortion = c(1007,1118,1032,4150,2533, 3201, 2533, 2812),
Sum = c(5096,9068,8412,9035,9589, 11158, 9231, 7344)
)
# 1. Linear regression model for Extorsion
model_extortion <- lm(Extortion ~ Year, data=AF)
```

```
# 2. Linear regression model for Malware
model_malware <- lm(Malware ~ Year, data=AF)
# 3. Linear regression model for Unauthorised_access
model_unauth_access <- lm(Unauthorised_access ~ Year, data=AF)
# Predictions for a given year (e.g., 2023)
year_data <- data.frame(Year = 2023,2024)
predicted_extortion <- predict(model_extortion, year_data)
predicted_malware <- predict(model_malware, year_data)
predicted_unauth_access <- predict(model_unauth_access, year_data)
# Printing the predictions for the year 2023
cat("Predicted Extorsion for 2023:", predicted_extortion, "\n")

Predicted Extorsion for 2023: 3616.107

cat("Predicted Malware for 2023:", predicted_malware, "\n")

Predicted Malware for 2023: 6196.929

cat("Predicted Unauthorised_access for 2023:", predicted_unauth_access, "\n")

Predicted Unauthorised_access for 2023: 161.25

# Build the linear regression model
model_sum <- lm(Sum ~ Year, data=AF)
# Summary of the model
summary(model_sum)


Call:
lm(formula = Sum ~ Year, data = AF)


Residuals:
    Min      1Q  Median      3Q     Max 
-2464.7  -687.1   408.6   917.5  2088.8 


Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -600369.6   545332.8  -1.101    0.313
Year            301.7      270.2   1.117    0.307
```

```
Residual standard error: 1751 on 6 degrees of freedom
Multiple R-squared:  0.1721,    Adjusted R-squared:  0.03409
F-statistic: 1.247 on 1 and 6 DF,  p-value: 0.3068

# Predictions for future years, for example 2023 and 2024
future_years <- data.frame(Year=c(2023, 2024))
predictions <- predict(model_sum, newdata=future_years)
print(predictions)

       1         2
 9974.286 10275.988

# Plot the original data and the linear regression line
ggplot(AF, aes(x=Year, y=Sum)) +
geom_point() +
geom_smooth(method="lm", se=FALSE, color="blue") +
labs(title="Year vs Sum Prediction from Action Fraud", x="Year", y="Sum") +
geom_text(aes(label=Sum), hjust=1.5, vjust=0.5)

`geom_smooth()` using formula = 'y ~ x'
```

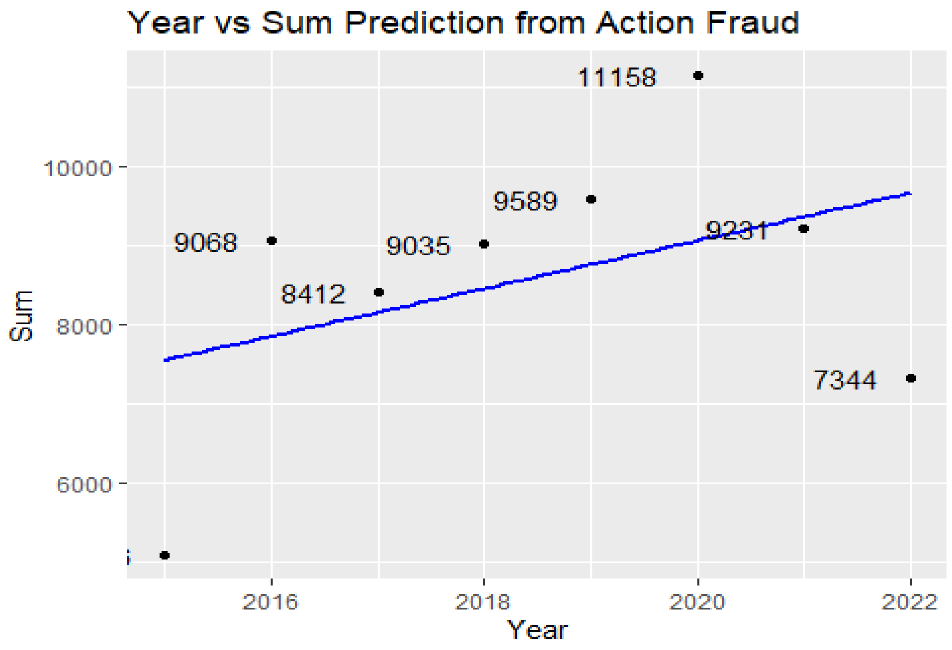


Fig 5: Regression model for Action fraud

```
#Plot graph for above prediction
# Create a data frame for predictions
predicted_data <- data.frame(
category = c("Extorsion", "Malware", "Unauthorised_access"),
predicted_value = c(3616, 6196,161)
)
# Combine original data and predictions
combined_data <- rbind(
AF[, c("Year", "Sum")],
data.frame(Year=future_years$Year, Sum=predictions)
)
# Plot using ggplot2
library(ggplot2)
p <- ggplot(combined_data, aes(x=Year, y=Sum)) +
geom_point(aes(color=ifelse(Year <= 2022, "Actual", "Predicted")), size=3) +
geom_line(aes(group=1)) +
scale_color_manual(values=c("Actual"="blue", "Predicted"="red")) +
labs(title="Year vs Sum Prediction from Action Fraud", x="Year", y="Sum", color="Data Type") +
geom_text(aes(label=round(Sum, 0)), vjust=-0.5, hjust=0.5) +
theme_minimal()

print(p)
```

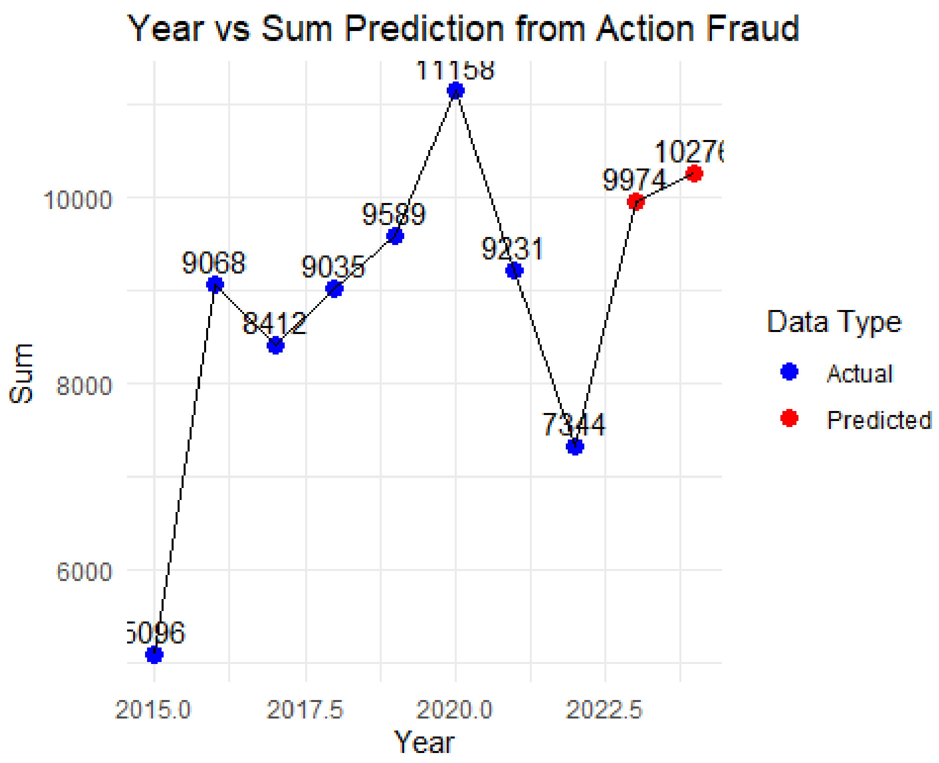


Fig 6: Prediction graph for Action fraud

A dataset named AF is constructed, containing data for the years 2015 through 2022. This dataset encompasses several types of cyber incidents: Malware, Unauthorised Access, Extortion, and a cumulative sum of all incidents (Sum).

The dataset serves as the foundation for creating linear regression models for each type of cyber incident.

Construction of Linear Regression Models:

The regression line is plotted closest to the data points in a regression graph.

Three separate linear regression models are constructed to predict the number of incidents for each cybercrime category based on the year:

1. **model_extortion**: Predicts incidents of extortion.
2. **model_malware**: Predicts incidents of malware.
3. **model_unauth_access**: Predicts incidents of unauthorised access.

Each model uses the year as an independent variable and the respective incident count as the dependent variable. The goal is to determine if there's a linear relationship between the year and the number of incidents.

Making Predictions:

Predictions are made for the years 2023 and 2024 using the linear regression models. The intent is to forecast the number of incidents for each category in these future years.

3.6.2 Regression model for ICO:

```
ICO <- data.frame(
Year <- c(2019, 2020, 2021, 2022),
Malware <- c(75, 101, 134, 70),
Other_cyber_incidents <- c(53, 213, 260, 232),
Ransomware <- c(77, 160, 529, 542),
Unauthorised_access <- c(473, 597, 619, 175),
Sum <- c(678, 1071, 1542, 1019)
)
# 1. Linear regression model for Ransomware
model_ransomware <- lm(Ransomware ~ Year, data=ICO)
# 2. Linear regression model for Malware
model_malware <- lm(Malware ~ Year, data=ICO)
# 3. Linear regression model for Unauthorised_access
model_unauth_access <- lm(Unauthorised_access ~ Year, data=ICO)
# 4. Linear regression model for Other_cyber_incidents
model_other_cyber <- lm(Other_cyber_incidents ~ Year, data=ICO)
# Predictions for a given year (e.g., 2023 and 2024)
year_data <- data.frame(Year = c(2023, 2024))
predicted_ransomware <- predict(model_ransomware, year_data)
predicted_malware <- predict(model_malware, year_data)
predicted_unauth_access <- predict(model_unauth_access, year_data)
predicted_other_cyber <- predict(model_other_cyber, year_data)
# Printing the predictions for the year 2023 and 2024
cat("Predicted Ransomware for 2023:", predicted_ransomware[1], "\n")

Predicted Ransomware for 2023: 768

cat("Predicted Malware for 2023:", predicted_malware[1], "\n")

Predicted Malware for 2023: 99.5

cat("Predicted Unauthorised_access for 2023:", predicted_unauth_access[1], "\n")

Predicted Unauthorised_access for 2023: 248
```

```
cat("Predicted Other_cyber_incidents for 2023:", predicted_other_cyber[1], "\n")

Predicted Other_cyber_incidents for 2023: 335.5

# Build the linear regression model for Sum
model_sum <- lm(Sum ~ Year, data=ICO)
# Summary of the model
summary(model_sum)


Call:
lm(formula = Sum ~ Year, data = ICO)


Residuals:
     1      2      3      4 
-175.4   68.2  389.8 -282.6 


Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -300785.2   330290.4  -0.911    0.459
Year            149.4      163.5   0.914    0.457


Residual standard error: 365.5 on 2 degrees of freedom
Multiple R-squared:  0.2946,    Adjusted R-squared:  -0.0581 
F-statistic: 0.8353 on 1 and 2 DF,  p-value: 0.4572

# Predictions for future years, 2023 and 2024
predictions <- predict(model_sum, newdata=year_data)
print(predictions)

     1      2 
1451.0 1600.4 

# Plot the original data and the linear regression line
ggplot(ICO, aes(x=Year, y=Sum)) +
geom_point() +
geom_smooth(method="lm", se=FALSE, color="blue") +
labs(title="Year vs Sum Prediction from ICO", x="Year", y="Sum") +
geom_text(aes(label=Sum), hjust=1.5, vjust=0.5)
```

```
`geom_smooth()` using formula = 'y ~ x'
```

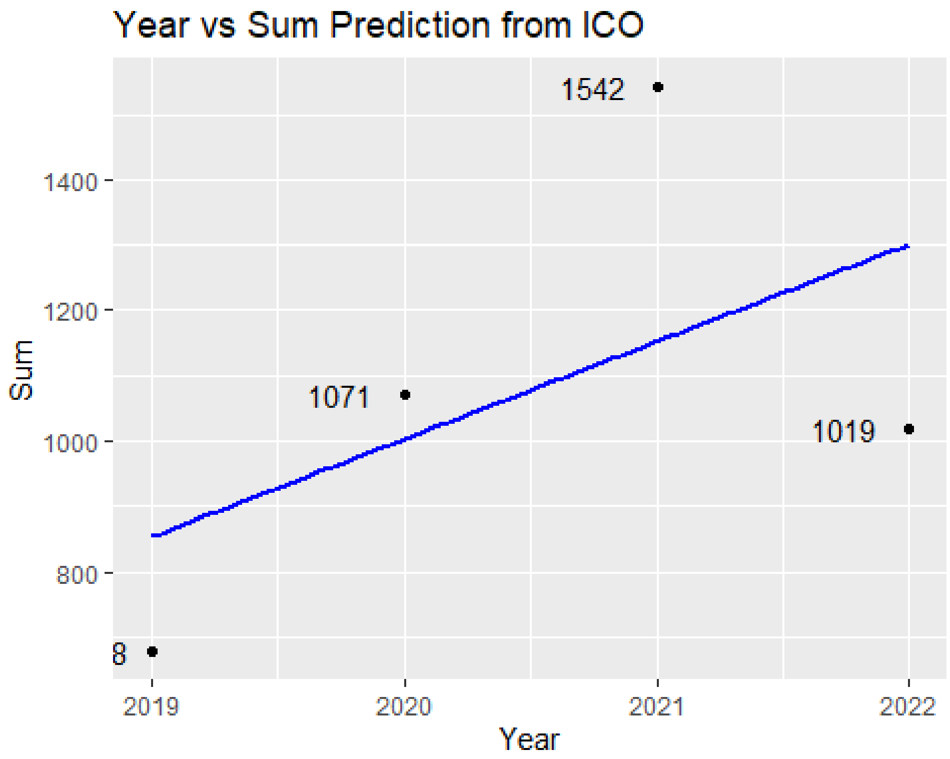


Fig 7: Regression model for ICO

```
# Create a data frame for predictions
predicted_data <- data.frame(
Year = c(2023, 2024),
Sum = predictions
)

# Creating ICO dataset
ICO <- data.frame(
  Year = c(2019, 2020, 2021, 2022),
  Malware = c(75, 101, 134, 70),
  Other_cyber_incidents = c(53, 213, 260, 232),
  Ransomware = c(77, 160, 529, 542),
  Unauthorised_access = c(473, 597, 619, 175),
  Sum = c(678, 1071, 1542, 1019)
)
# Linear regression models
model_sum <- lm(Sum ~ Year, data=ICO)
```

```r
# Predictions for 2023 and 2024
year_data <- data.frame(Year = c(2023, 2024))
predictions <- predict(model_sum, newdata=year_data)

# Combining original and predicted data
combined_data <- rbind(
  ICO[, c("Year", "Sum")],
  data.frame(Year=year_data$Year, Sum=predictions)
)

# Plotting the data
p <- ggplot(combined_data, aes(x=Year, y=Sum)) +
  geom_point(aes(color=ifelse(Year <= 2022, "Actual", "Predicted")), size=3) +
  geom_line(aes(group=1)) +
  scale_color_manual(values=c("Actual"="blue", "Predicted"="red")) +
  labs(title="Year vs Sum Prediction from ICO", x="Year", y="Sum", color="Data Type") +
  geom_text(aes(label=round(Sum, 0)), vjust=-0.5, hjust=0.5) +
  theme_minimal()

print(p)
```

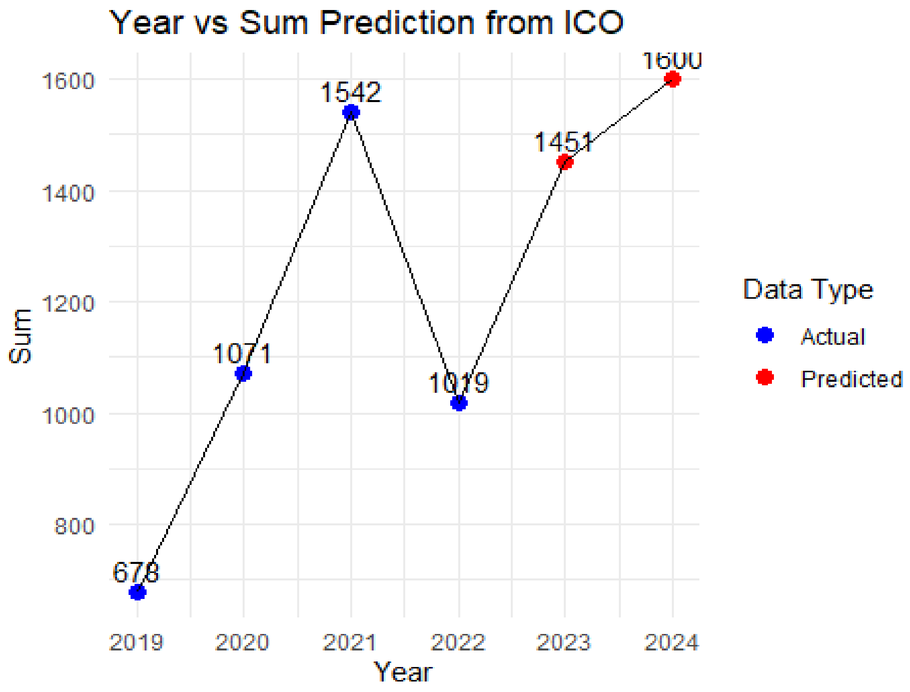


Fig 8: Prediction graph for ICO

A data frame named ICO is created which contains data for multiple years (2019 to 2022). The columns in this data frame represent different types of cyber incidents, namely:

1. Malware
2. Other cyber incidents
3. Ransomware
4. Unauthorised access
5. The last column, Sum, aggregates the number of incidents for each year.

Linear Regression Models:

Four separate linear regression models are constructed to predict the number of incidents for each cyber incident type based on the year. The models are:

model_ransomware for predicting ransomware incidents.

model_malware for predicting malware incidents.

model_unauth_access for predicting unauthorised access incidents.

model_other_cyber for predicting other types of cyber incidents.

Each model uses the year as an independent variable and the count of a particular type of cyber incident as the dependent variable. The goal is to determine if there's a linear trend in the number of incidents over the years.

Predictions:

1. Predictions are made for the years 2023 and 2024 using the constructed linear regression models. This is done to forecast the number of incidents in these years for each cyber incident type.
2. The predictions are stored in variables like predicted_ransomware, predicted_malware, predicted_unauth_access, and predicted_other_cyber.

4. Results

4.1 Findings

1. **What association or correlation is there between the Action Fraud and ICO Cybercrime Data?**

Malware Incidents:

The number of malware incidents reported by Action Fraud (AF) is noticeably higher than those reported by the Information Commissioner's Office (ICO). When tested this difference statistically, it showed that there might be a significant difference.

The result of the statistical test indicates that there might be a significant difference in the number of malware incidents reported by AF and ICO. This implies that the disparity between the reported malware incidents from the two organisations is not likely to be due to random chance but could be attributed to other factors.

Unauthorised Access Incidents:

In contrast, the number of malware incidents reported by Action Fraud (AF) is noticeably higher than those reported by the Information Commissioner's Office (ICO). When tested this difference statistically, it showed that there might be a significant difference.

The outcome of the statistical test suggests that the difference in the number of unauthorised access incidents reported by AF and ICO might not be significant or consistent. The variation observed could potentially be due to random fluctuations rather than substantial differences.

2. **Is there sufficient evidence that successful ICO enforcement action and prosecution can be associated or correlated with cyber security breaches?**

To address the main question directly, if only have correlation values related to years and not direct data on successful ICO enforcement actions or prosecutions, we can’t make a definitive statement. The correlations do hint at patterns, but they don’t directly indicate a correlation between successful ICO enforcement and the occurrence of cyber security breaches. To provide a concrete answer, we would need to calculate the correlation between ICO enforcement actions/prosecutions and actual cyber security breach incidents. Furthermore, other external factors and context would also need to be considered. Based on the provided data, while there’s evidence of trends in reporting to ICO and Action Fraud, there isn't direct evidence to conclusively say that successful ICO enforcement action and prosecution are associated or correlated with cyber security breaches. Additional data such as out of all the crimes reported to the Action fraud, how many crimes had law enforcement on them, and analysis would be necessary to draw a more definitive conclusion.

An illustrative example of how the dataset should be created to get the above analysis:

Year	Number_of_Breaches	ICO_Enforcement_Actions	Prosecutions
2018	100	5	4
2019	110	8	7
2020	150	15	12
2021	160	20	18

In this dataset: - Year refers to the year the data was recorded. - Number_of_Breaches refers to the total number of reported cybersecurity breaches in that year. - ICO_Enforcement_Actions refers to the number of enforcement actions taken by the Information Commissioner's Office (ICO) that year in response to breaches. - Prosecutions refers to the number of legal prosecutions that resulted from those breaches.

Analysis: If such dataset exists, we can look if there is a trend indicating that as cybersecurity breaches increase, there is a proportional increase in ICO enforcement actions and prosecutions? Statistical tests, like correlation analysis, would provide quantitative measures of any associations.

If there's a strong positive correlation between the Number_of_Breaches and ICO_Enforcement_Actions or Prosecutions, it might suggest that there is indeed an association.

3. **To what extent can future cybercrime and fraud be predicted using the available historic data?**

Action Fraud's 2023 Forecasts:

a. Extortion: ~3,616

b. Malware: ~6,197

c. Unauthorised Access: ~161

d. Total incidents for 2023: 9,974 and for 2024: 10,276.

ICO’s 2023 Forecasts:

a. Ransomware: ~768
b. Malware: ~99.5
c. Unauthorised Access: ~248
d. Other Cyber Incidents: ~335.5
e. Total incidents for 2023: 1,451 and for 2024: 1,600

To find a trend line from previous data, linear regression to forecast future is used. Action Fraud forecasts an increase of about 302 occurrences each year, whereas ICO predicts an increase of about 150.

Testing both models on historical data revealed their flaws: - Only 17% of historical fluctuations were accounted for by Action Fraud's model. After corrections, this fell further to 3.41%. - The ICO model initially accounted for 29% of the historical volatility. A closer look reveales overfitting. - With p-values above the 0.05 cut off, statistical tests, such as the F-statistic and p-values, suggested that trends might be more the result of random fluctuations than actual patterns.

Predictions shown graphically in red, while actual data was shown in blue. While some locations aligned, significant variations exposed the models' limitations.

Because of the constantly shifting nature of the digital world, which is influenced by technology, policy, major world events, etc., predicting cybercrime is difficult. The model's obvious flaws were on display. Historical data can offer insights, but it might not be the best tool for prediction given the changing nature of cyber threats.

5.Discussion and recommendations

1.1 Comparison of cyber-crime Organisations in India Vs UK

1. **Indian Computer Emergency Response Team (CERT-In):** Established by the Department of Information Technology, CERT-In is the national agency for responding to computer security incidents. It provides alerts and advisories to mitigate cyber threats (Chaturvedi et al., 2008).
2. **National Critical Information Infrastructure Protection Centre (NCIIPC**): This organisation is dedicated to the protection of critical information infrastructure in India, safeguarding sectors like power, transport, and telecommunications from cyber threats (Nickolov, 2005).
3. **Cyber-Crime Cells:** Major cities in India have cybercrime cells that deal with online offenses such as fraud, defamation, and data theft. They are often the first point of contact for individuals and organisations facing cyber threats (Singh et al., 2023).

Scope & Jurisdiction:

Action Fraud and ICO have specific jurisdictions, with the former focusing on reporting fraud and the latter on data protection. In contrast, India's agencies have broader mandates. CERT-In handles a wide range of cyber threats, NCIIPC focuses on critical infrastructure, and Cyber-Crime Cells address localized cybercrimes (Chaturvedi et al., 2008).

Public Interaction:

Action Fraud provides a direct platform for the UK public to report cybercrimes. Similarly, various states' Cyber-Crime Cells in India allow the public to report cyber offenses directly. CERT-In and NCIIPC, on the other hand, primarily work behind the scenes, collaborating with organisations and providing them with critical cybersecurity information (Bada et al., 2019).

Regulatory Power:

The ICO has considerable regulatory powers, including the ability to levy hefty fines on organisations that violate data protection laws. In contrast, while agencies like CERT-In have advisory roles, they do not possess the same punitive powers as the ICO (Chaturvedi et al., 2008).

Protective vs. Reactive Approach:

NCIIPC and CERT-In often adopt a proactive approach, identifying potential threats and vulnerabilities and issuing advisories. On the other hand, Action Fraud and Cyber-Crime Cells typically react to incidents reported to them, taking necessary actions post-incident (Singh et al., 2023).

Collaboration & Partnerships:

Both UK and Indian agencies collaborate with international counterparts. For instance, ICO works alongside European data protection agencies, while CERT-In collaborates with its counterparts in various countries to share threat intelligence (Singh et al., 2023).

5.2 Data Harmonization

Cybercrimes are becoming a bigger threat in today's rapidly changing digital world. The obvious inconsistencies in their stated cybercrime statistics, however, is a concerning problem. These variations may cause stakeholders and policymakers to form inaccurate perceptions when attempting to determine the trends in cybercrime.

There is an urgent necessity for integrating the data supplied by these organisations, highlighting the significance of data integrity. Data harmonization is the process of assuring compatibility and uniformity among datasets, regardless of where they came from. This relates to establishing how businesses like AF and ICO represent various cybercrime metrics like malware attacks or unauthorised access.

This primarily enables strategic allocation of resources, enabling policymakers to react to current cyber threats more successfully. Additionally, data consistency from two respected sources increases public trust. When published data from different bodies agree, the general public and stakeholders are more likely to believe the data.

The cooperation of AF and ICO offers a potential remedy for closing this data gap. Both organisations can effectively pool, analyze, and evaluate their data by creating a joint platform or framework. Such collaborative actions not only strengthen the reported data's accuracy but also strengthen its credibility by repeating shared conclusions. In essence, the AF and ICO can present a more cohesive front against cybercrimes through strategic partnership and data standardization.

5.3 Focused Efforts on Malware

Malware stands out as an especially common worry in the vast array of dangers that make up the cyber world. A substantial discrepancy is seen between the data from Action Fraud (AF) and the Information Commissioner's Office (ICO): AF records a significantly larger number of malware instances than ICO. The difference is more than just a number; it raises questions about possible knowledge, reporting, or both gaps among platform users.

Given the previously mentioned gap, it is essential that AF step up its user education and malware defence efforts. It may be assumed that users of AF are either more frequently targeted or more diligent in reporting malware because of the greater occurrence of malware reports with AF. Increasing user education can be a successful preventive in either circumstance. AF may actively lessen the vulnerability of its user base by making sure people are aware of the signs of malware assaults and the precautions to take.

However, AF shouldn't adopt a unique strategy. Collaboration may be the key to increased effectiveness. ICO has a wealth of knowledge and tactics while having fewer instances of reported malware. A collaborative strategy in which AF and ICO share information, strategies, and best practises could be advantageous to both organisations. By combining the expertise and assets of the two organisations, such cooperative efforts can promote a unified front against cyber-attacks.

5.4 Improve Unauthorised Access Reporting

Unauthorised access events recorded by ICO are more frequent than those by AF. Both organisations should assess the causes of this difference. They should investigate the reasons behind any underreporting of specific unlawful accesses in one dataset.

It is critical that AF and ICO look more closely at the underlying factors causing this variation in their datasets. An initial suggestion would be for both organisations to do a thorough examination and analysis of their own data-gathering processes, reporting processes, and user engagement programmes. Are there any structural or operational distinctions that favour one organisation over the other in terms of making it simpler for victims to report unauthorised access? Or does one organisation define unauthorised access more broadly to include a larger range of incidents?

Furthermore, a key focus should be on the potential for underreporting in either dataset. Underreporting can occur for a variety of reasons, including user ignorance, suspicion about the reporting system's capacity to generate actionable results, or even fear of possible consequences. Both organisations can customise their outreach and teaching efforts to close these gaps by determining the causes of underreporting. For instance, public awareness campaigns and educational measures could be stepped up if a lack of understanding is found to be a major contributing factor.

Furthermore, AF and ICO need to think about working together and exchanging ideas and best practises. By working together, both parties can improve their methods and ensure a more thorough and coordinated response to situations involving unauthorised access.

5.5 Enhance Predictive Modelling

As the digital environment changes, so do the more complex cyberthreats. In order to predict and respond to possible cyber issues, Action Fraud (AF) and the Information Commissioner's Office (ICO) extensively rely on predictive modelling. However, the present prediction models used by both AF and ICO have several flaws that could be reducing their effectiveness.

There is an immediate requirement for AF and ICO to reconsider their predictive modelling strategy in light of these limitations. Relying only on conventional prediction approaches may not be sufficient as cyber threats become more complex. The development of advanced predictive analytics technologies, particularly those driven by machine learning (ML) and artificial intelligence (AI), presents a promising way to improve the accuracy of their projections.

Large and varied datasets can be analysed by AI and ML, which can also identify patterns that are frequently invisible to human analysts and quickly adjust to new data. Both organisations can be more prepared and more proactive thanks to these skills, which can offer more precise threat forecasts. AF and ICO can develop deeper comprehension of emerging threats and more focused reaction plans by using AI and ML into their predictive models.

Utilising AI and ML can also result in faster response times, real-time threat identification, and perhaps a considerable decrease in successful cyberattacks. The advantage of continuous learning provided by these technologies allows the models to improve and adapt their predictions as new data becomes available, ensuring that they continue to be accurate and useful over time.

In conclusion, it's critical for AF and ICO to use cutting-edge predictive technology if they want to stay one step ahead of cyber threats. The incorporation of AI and ML into their predictive frameworks can significantly improve the accuracy of their forecasts, resulting in more efficient and prompt responses to cyber threats.

5.6 Public Awareness

Developing a well-informed public that can identify and report threats is also an important part of an efficient response process. The differences between the cyber incident statistics provided by Action Fraud (AF) and the Information Commissioner's Office (ICO) indicate different public knowledge and participation levels with these two platforms.

The general public must be aware of both the dangers they confront and the locations and procedures for reporting potential threats and security breaches if we are to maintain the overall safety and security of the digital environment. The gap in numbers between AF and ICO may indicate that the public favours one platform over the other or may not be fully aware of both platform’s capabilities.

Both organisations must work together in the area of public relations to address this potential knowledge gap. The AF and ICO can encourage more informed and proactive digital citizens by increasing their visibility and outreach. The goal here is to inform the public about the platform’s existence as well as the value of reporting, the procedures involved, and the consequences.

Online resources, community involvement, workshops, and collaborative campaigns can all be helpful. Also helpful in clarifying the reporting process and highlighting its significance are monthly updates on the state of cyber threats, success stories resulting from public reporting, and clear, user-friendly rules.

However, the public's engagement is crucial for an integrated approach to cybersecurity. Therefore, it is essential that AF and ICO step up their public relations efforts. Both organisations may use collective vigilance to build a safer online environment for everyone by making sure the public is informed and involved.

5.7 Collaboration with Tech Industry

In an age of fast technological development, the variety of cyber threats is always changing. It is challenging not only to identify these threats, but also to predict and deal with newly emerging ones. The Information Commissioner's Office (ICO) and Action Fraud (AF), two significant bodies addressing cyber issues, play critical roles. However, it is clear that their efforts would benefit from stronger collaboration with the IT industry.

5.8 Feedback Mechanism

Create a feedback mechanism using which the two organisations can work together to improve their respective data collection and analysis techniques by exchanging insights from their datasets.

The public's trust will rise because of constant reporting and a collaborative effort to address cyberthreats, which will also improve the efficiency of counter-cybercrime efforts. Collaboration, improved modelling, and more public knowledge are critical in this regard.

6. Conclusion

Insights into the world of cyber threats can be found through the examination of cybercrime data from Action Fraud (AF) and the Information Commissioner's Office (ICO). The two authorised organisations inconsistencies in data may provide inaccurate impressions, which may affect how resources are allocated and how policies are developed. Although predictive modelling is widely used to predict possible cyber risks, these models still have space for development in terms of performance.

While some may consider cybercriminal and security dangers as an extension of existing threats to business, the threat landscape is an extremely unpredictable one. Only seven years ago, some criminologists cautioned that 'those who fail to anticipate the future are in for an unpleasant shock when it arrives (Smith et al., 2004). It is critical for our society to be ready, as well as for businesses and organisations to develop faster than criminals and other bad factors.

6.2 Further research

The foundation of any robust analysis lies in the quality and comprehensiveness of the dataset at hand. A pivotal direction for future research entails the creation of a dataset meticulously curated with pertinent attributes related to cyber threats. Such a dataset would serve as the bedrock for exhaustive analyses, helping decode the myriad dimensions of cyber threats and their implications. Given the unpredictable and dynamic nature of the cyber threat landscape, a critical area of exploration is discerning strategies that are not merely reactive but inherently proactive. Research could delve into technological, organisational, and strategic avenues that enable entities to not only respond to threats but also to anticipate them.

Beyond technological defences, the organisational ethos plays a pivotal role in cybersecurity. An essential research trajectory involves understanding how institutions can embed a culture of continuous learning and adaptability. Such a culture would ensure that they remain agile and resilient, capable of navigating the flux of the digital realm and its associated threats. A comprehensive examination of the prevailing legislative and regulatory frameworks is crucial. Research should assess the deterrent power of these frameworks against cybercrimes and their efficacy in ensuring accountability. Such an analysis would provide insights into potential gaps and areas of enhancement. Building upon the assessment of existing legal frameworks, future research should delve into potential modifications or augmentations that could bolster cybersecurity efforts. Exploring how legal paradigms can be fine-tuned to enhance prevention, facilitate timely detection, and ensure rigorous prosecution of cybercrimes would be of paramount importance. It is essential to discern strategies that offer maximum protection. Research could focus on identifying and evaluating preventive measures tailored for businesses and organisations, ensuring they are insulated from the multifaceted risks posed by cyber threats.

Adopting a global lens, it is imperative to explore international best practices in countering cyber threats. Such research would entail a comparative analysis of strategies adopted by different nations, discerning patterns of success, and understanding how these can be integrated into the national cybersecurity strategy. Understanding these global paradigms would not only offer insights into effective strategies but also foster international collaboration in the collective endeavour against cyber threats.

References

Aiken, M., 2016. *The cyber effect: A pioneering cyber-psychologist explains how human*

Akdemir, N. and Lawless, C.J., 2020. Exploring the human factor in cyber-enabled and cyber-dependent crime victimisation: A lifestyle routine activities approach. Internet Research, 30(6), pp.1665-1687.

Ali, A. (2017) *Ransomware: A research and a personal case study of dealing with this ..., Information Systems*. Available at: https://www.iup.edu/business/files/for_faculty_and_staff/azad-ali-2017-publication_2018-award_ransomware.pdf.

Alotaibi, M., Furnell, S. and Clarke, N. (2016) 'Information security policies: A review of challenges and influencing factors', *2016 11th International Conference for Internet Technology and Secured Transactions (ICITST)* [Preprint]. doi:10.1109/icitst.2016.7856729.

Al-rimy, B.A., Maarof, M.A. and Shaid, S.Z. (2018) 'Ransomware threat success factors, taxonomy, and countermeasures: A survey and research directions', *Computers &amp; Security*, 74, pp. 144–166. doi:10.1016/j.cose.2018.01.001.

Anderson, R. and Moore, T. (2006) 'The Economics of Information Security', *Science*, 314(5799), pp. 610–613. doi:10.1126/science.1130992.

Bada, M., Sasse, A.M. and Nurse, J.R., 2019. Cyber security awareness campaigns: Why do they fail to change behaviour?. *arXiv preprint arXiv:1901.02672*.

Bada, M., Sasse, A.M. and Nurse, J.R.C. (2019) *Cyber Security Awareness Campaigns: Why do they fail to change behaviour?* [Preprint]. doi: https://doi.org/10.48550/arXiv.1901.02672.

Bendovschi, A. (2015) 'Cyber-attacks – trends, patterns and security countermeasures', *Procedia Economics and Finance*, 28, pp. 24–31. doi:10.1016/s2212-5671(15)01077-1.

Bhaimia, S. (2018) ‘The General Data Protection Regulation: THE NEXT GENERATION OF EU Data Protection’, *Legal Information Management*, 18(1), pp. 21–28. doi:10.1017/s1472669618000051.

Broadhurst, R., Grabosky, P., Alazab, M., Bouhours, B. and Chon, S., 2014. An analysis of the nature of groups engaged in cyber-crime. *An analysis of the nature of groups engaged in cyber-crime, International Journal of Cyber Criminology*, *8*(1), pp.1-20.

Buckley, J. (2021) ‘The industrialisation of cyber extortion’, *Computer Fraud &amp; Security*, 2021(12), pp. 6–10. doi:10.1016/s1361-3723(21)00127-5.

Buil-Gil, D., Miró-Llinares, F., Moneva, A., Kemp, S. and Díaz-Castaño, N., 2021. Cybercrime and shifts in opportunities during COVID-19: a preliminary analysis in the UK. European Societies, 23(sup1), pp.S47-S59.

Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., & Chandra, T. (2016). Bigtable: A distributed storage system for structured data. ACM Transactions on Computer Systems (TOCS), 26(2), 1-26.

Chaturvedi, M., Gupta, M. and Bhattacharya, J. (2008) *Cyber Security Infrastructure in India: A study - researchgate*, *CSI*. Available at: https://www.researchgate.net/publication/228846974_Cyber_Security_Infrastructure_in_India_A_Study (Accessed: 01 August 2023).

Chen, Q. and Bridges, R.A. (2017) ‘Automated Behavioral Analysis of Malware: A case study of WannaCry ransomware’, *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)* [Preprint]. doi:10.1109/icmla.2017.0-119.

Clough, J. (2015) *Principles of cybercrime*. Cambridge: Cambridge University Press.

Collier, B., Horgan, S., Jones, R. and Shepherd, L., 2020. The implications of the COVID-19 pandemic for cybercrime policing in Scotland: A rapid review of the evidence and future considerations. Scottish Institute for Policing Research, available at: https://www.sipr.ac.uk/wp-content/uploads/2021/10/The-implications-of-the-COVID-19-pandemic-for-cybercrime-policing-in-Scotland-A-rapid-review-of-the-evidence-and-future-considerations.pdf,

Crowther, G.A., 2017. The cyber domain. The cyber defense review, 2(3), pp.63-78.

Ell, M. and Gallucci, R. 2022. Cyber Security Breaches Survey 2022, available at: https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2022/cyber-security-breaches-survey-2022,

Emigh, A. (2006) 'The crimeware landscape: Malware, phishing, identity theft and beyond', *Journal of Digital Forensic Practice*, 1(3), pp. 245–260. doi:10.1080/15567280601049985.

FBI (2016) *Ransomware prevention and response for Cisos, FBI*. Available at: https://www.fbi.gov/file-repository/ransomware-prevention-and-response-for-cisos.pdf/view.

Furnell, S.M. and Warren, M.J. (1999) 'Computer Hacking and cyber terrorism: The real threats in the new millennium?', *Computers &amp; Security*, 18(1), pp. 28–34. doi:10.1016/s0167-4048(99)80006-6.

Giro Correia, S. (2022) 'Making the most of cybercrime and fraud crime report data: A case study of uk action fraud', *International Journal of Population Data Science*, 7(1). doi:10.23889/ijpds.v7i1.1721.

Harit, A., Ezzati, A. and Elharti, R. (2017) 'Internet of things security', *Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing* [Preprint]. doi:10.1145/3018896.3056784.

HM Government, 2016. National Cyber Strategy 2016 to 2021, available at: https://www.gov.uk/government/publications/national-cyber-security-strategy-2016-to-2021,

HM Government, 2022. National Cyber Strategy 2022, available at: https://www.gov.uk/government/publications/national-cyber-strategy-2022,

Horgan, S., Collier, B., Jones, R. and Shepherd, L., 2021. Re-territorialising the policing of cybercrime in the post-COVID-19 era: towards a new vision of local democratic cyber policing. Journal of Criminal Psychology, 11(3), pp.222-239.

Houser, W. (2015) 'Could what happened to Sony happen to us?', *IT Professional*, 17(2), pp. 54–57. doi:10.1109/mitp.2015.21.

Iriberri, A. and Navarrete, C.J. (2010) ‘Internet crime reporting: Evaluation of a crime reporting and investigative interview system by comparison with a non-interactive reporting alternative’, *2010 43rd Hawaii International Conference on System Sciences* [Preprint]. doi:10.1109/hicss.2010.460.

Jagatic, T.N. *et al.* (2007) ‘Social phishing’, *Communications of the ACM*, 50(10), pp. 94–100. doi:10.1145/1290958.1290968.

Johns, E. and Ell, M. 2023a. Cyber Security Breaches Survey 2023, available at: https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2022,

Johns, E. and Ell, M. 2023b. Cyber Breaches Survey 2023: education institutions annex, available at: https://www.gov.uk/government/statistics/cyber-security-breaches-survey-2023/cyber-security-breaches-survey-2023-education-institutions-annex,

Kemp, S., 2023. Exploring public cybercrime prevention campaigns and victimization of businesses: A Bayesian model averaging approach. Computers & Security, 127, p.103089.

Kharaz, A. *et al.* (1970) *{unveil}: A {large-scale}, automated approach to detecting ransomware*, *USENIX*. Available at: https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/kharaz.

Kianpour, M., Kowalski, S.J. and Øverby, H., 2021. Systematically understanding cybersecurity economics: A survey. *Sustainability*, *13*(24), p.13677.

Kumaraguru, P. *et al.* (2007) ‘Protecting people from phishing’, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* [Preprint]. doi:10.1145/1240624.1240760.

Lastdrager, E.E. (2014) ‘Achieving a consensual definition of phishing based on a systematic review of the literature’, *Crime Science*, 3(1). doi:10.1186/s40163-014-0009-y.

Levi, M., 2017. Assessing the trends, scale and nature of economic cybercrimes: overview and Issues: In Cybercrimes, cybercriminals and their policing, in crime, law and social change. Crime, law and social change, 67, pp.3-20.

Linden, T. *et al.* (2019) *The privacy policy landscape after the GDPR*, *arXiv.org*. Available at: https://arxiv.org/abs/1809.08396 (Accessed: 02 August 2023).

MAIMON, D. *et al.* (2013a) 'Restrictive deterrent effects of a warning banner in an attacked Computer System', *Criminology*, 52(1), pp. 33–59. doi:10.1111/1745-9125.12028.

Maniath, S., Poornachandran, P. and Sujadevi, V.G. (2019) 'Survey on prevention, mitigation and containment of ransomware attacks', *Communications in Computer and Information Science*, pp. 39–52. doi:10.1007/978-981-13-5826-5_3.

McGuire, M. and Dowling, S., 2013. Chapter 2: Cyber-enabled crimes-fraud and theft. *Cyber-crime: A review of the evidence*.

McGuire, M., 2007. *Hypercrime: The new geometry of harm*. Routledge.

National Cyber Security Centre 2019. Incident Management; Plan: Your cyber incident response processes, available at: https://www.ncsc.gov.uk/collection/incident-management/cyber-incident-response-processes,
National Cyber Security Centre 2021. 10 Steps to Cyber Security; Risk Management, available at: https://www.ncsc.gov.uk/collection/10-steps/risk-management,

Nickolov, E. (2005) *Critical Information Infrastructure Protection: Analysis, evaluation.*, *ResearchGate*. Available at: https://www.comw.org/tct/fulltext/05nickolov.pdf
(Accessed: 02 August 2023).

Parliament, U. (1990) *Computer misuse act 1990*, *Legislation.gov.uk*. Available at: https://www.legislation.gov.uk/ukpga/1990/18/contents (Accessed: 01 August 2023).

Radanliev, P. *et al.* (2018) 'Economic impact of IOT cyber risk - analysing past and present to predict the future developments in IOT risk analysis and IOT Cyber Insurance', *Living in the Internet of Things: Cybersecurity of the IoT - 2018* [Preprint]. doi:10.1049/cp.2018.0003.

Rai, A.K., Tewari, R.R. and Upadhyay, S.K. (2010) *Different Types of Attacks on Integrated MANET-Internet Communication*. Available at:

https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=2f5695a5b6e6cea30cf aa817e960f5bca503cb9e.

Rai, A.K., Tewari, R.R. and Upadhyay, S.K., 2010. Different types of attacks on integrated manet-internet communication. *International Journal of Computer Science and Security*, *4*(3), pp.265-274.

Robinson, N. and Graux, H. (2009) *Technical report: Review of the European Data Protection directive, Review of the European Data Protection Directive* . Available at: https://www.huntonak.com/files/webupload/PrivacyLaw_review_of_eu_dp_directive.pdf.

Sgandurra, D., Muñoz-González, L., Mohsen, R. and Lupu, E.C., 2016. Automated dynamic analysis of ransomware: Benefits, limitations and use for detection. *arXiv preprint arXiv:1609.03020*.

Singh, V., Malik, V. and Mittal, R. (2023) 'Challenges to cybercrime reporting, investigation, and adjudication in India', *Advancements in Cybercrime Investigation and Digital Forensics*, pp. 1–24. doi:10.1201/9781003369479-1.

Smith, R.G., 2000, August. Confronting Fraud in the Digital Age. In *Fraud Prevention and Control Conference, the Australian Institute of Criminology in association with the Commonwealth Attorney-General's Department, Surfers Paradise* (pp. 24-25).

Tavani, H. (2007) 'Ethics and technology: Ethical issues in an age of information and communication technology', *Journal of Information, Communication and Ethics in Society*, 5(1), pp. 43–44. doi:10.1108/14779960710822683.

Zhang, Y. *et al.* (2018) *Phinding Phish: Evaluating anti-phishing tools*, *figshare*. Available at: https://kilthub.cmu.edu/articles/journal_contribution/Phinding_Phish_Evaluating_Anti-Phishing_Tools/6470321.

Appendix

Appendix 1 – Research Proposal Formative Feedback

Norwich Business School **RESEARCH PROPOSAL** **MARKER'S INDIVIDUAL FORMATIVE FEEDBACK**						
Submission Date						
Student Number						
Dissertation Title						
	On a scale of 1 to 5, where 1 is POOR and 5 is EXCELLENT, rate the proposal on its presentation of the following criteria					
Marking Criteria	**1**	**2**	**3**	**4**	**5**	**Comments**
Originality of research topic						
Research objectives and questions						
Relevant literature						
Appropriateness of methodology and methods						
Plan for data collection and plan for data analysis						
Timeplan						

Referencing and presentation						
FEED FORWARD						

Appendix 2 - Record of Dissertation Supervision Sessions

This form should be completed by the student, and signed by the supervisor, at the end of every supervision session

Completed forms should be emailed to <**nbs.msc@uea.ac.uk**>

The purpose of this form is to encourage critical reflection by the student on the research and learning process, to facilitate communication between the student and her/his supervisory team and to ensure that progression can be more easily assessed.

Student Number: 100383119 Student Name: Tanvi Shridhar Shetty

Name of Supervisor: Stephen J. Jones **Signature of Supervisor:** SJones

Date of Meeting: **16/03/2023** **Date/Time of Next Meeting**: **12/07/2023**

Main Issues Discussed:

We discussed a variety of topics. Our focus was on generating research ideas, learning how to structure research, and developing effective writing skills. My supervisor provided valuable insights and guidance on each topic, which will help me have a productive research journey. This meeting was an important first step in preparing for the next phases of the project.

Course of Action for Next Meeting:

Review the concepts we discussed previously, determine the most outstanding ones, and begin outlining our investigation. In addition, I have initiated research on existing literature about our subject matter to identify where our study fits in and what aspects require further exploration.

University of East Anglia | NORWICH BUSINESS SCHOOL

Record of Dissertation Supervision Sessions

This form should be completed by the student, and signed by the supervisor, at the end of every supervision session.

Completed forms should be emailed to <**nbs.msc@uea.ac.uk**>

The purpose of this form is to encourage critical reflection by the student on the research and learning process, to facilitate communication between the student and her/his supervisory team and to ensure that progression can be more easily assessed.

Student Number: **100383119** **Student Name:** **Tanvi Shridhar Shetty**

Name of Supervisor: **Stephen J. Jones** **Signature of Supervisor:** SJones

Date of Meeting: **12/07/2023** **Date/Time of Next Meeting**: **01/08/2023**

Main Issues Discussed:

The meeting mainly discussed improving the initial research ideas and assessing a research proposal, which included its objectives and methodologies. Moreover, a preliminary analysis idea was shared to contextualize the proposed research in the academic field and highlight its distinct contributions. The conversations gave useful guidance for the research's upcoming actions.

Course of Action for Next Meeting:

we'll focus on reviewing the analysis of my research. I'll present the key points, and we'll discuss feedback and possible improvements. We'll check the structure and delve into specifics. Additionally, we'll explore any new findings to include and review our project timeline.

UEA University of East Anglia | NORWICH BUSINESS SCHOOL

Record of Dissertation Supervision Sessions

This form should be completed by the student, and signed by the supervisor, at the end of every supervision session.

Completed forms should be emailed to <**nbs.msc@uea.ac.uk**>

The purpose of this form is to encourage critical reflection by the student on the research and learning process, to facilitate communication between the student and her/his supervisory team and to ensure that progression can be more easily assessed.

Student Number: **100383119** **Student Name:** **Tanvi Shridhar Shetty**

Name of Supervisor: **Stephen J.Jones** **Signature of Supervisor:** SJones

Date of Meeting: **01/08/2023** **Date/Time of Next Meeting**: ______________________

Main Issues Discussed:

The analysis was carefully reviewed, focusing on its goals and methods. Suggestions were given to improve certain areas, and the overall structure was examined for consistency. Furthermore, we discussed the possibility of incorporating new discoveries and assessed the project's advancement according to the timeline.

Course of Action for Next Meeting:

Appendix 3 – Marking Form

Dissertation Marking Form

Student Number:		Module Code:	

Criteria	**Weight (%)**	**% Mark**	**Comments**
Originality & Understanding Is the addressed topic original and novel? Is there a good understanding of the chosen topic? Are the research objectives clear? Are the conclusions and implications clearly placed to the relevant literature? Does the analysis of key findings reveals sufficient understanding of the topic?	25		
Arguments and Criticality Are the relevant theories critically assessed? Are the results critically discussed (e.g. compared to existing literature)? Are the main research questions and aims sufficiently well-developed?	30		

Data collection and Analysis Is the dataset appropriate for the purpose of the dissertation? Is the research design appropriate? Is the description of the methodology clear? Is the discussion of results detailed and straightforward? Are the limitations of the chosen data and methods properly analysed?	30		
Presentation, Structure & Written Expression Is there a logical structure? Is the writing clear to understand and at academic standards? Is there enough evidence of reading (relevant citations in appropriate places)? Is there adherence to conventions and standards given, including the referencing method?	15		
Total			

Appendix 4 – MSc dissertation submission checklist

Student ID: 100383119

I confirm that I have included the following in my dissertation:

A declaration of my contribution to the work and its suitability for the degree	Yes
An abstract of the work completed	Yes
A table of contents	Yes
A list of figures & tables (if applicable)	Yes
A glossary of terms (where appropriate)	None
A full reference list in Harvard style	Yes
Supervision meeting records with supervisor signature- THREE	Yes

Date: 31 August 2023