2019

An Evaluation of the Information Security Awareness of University Students

Alan Pike
Technological University Dublin

An Evaluation of the Information Security Awareness of University Students

Alan Pike

D16124621

A dissertation submitted in partial fulfilment of the requirements of Technological University Dublin for the degree of M.Sc. in Computer Science (Security and Forensics)

2019

I Alan Pike, certify that this dissertation which I now submit for examination for the award of MSc in Computing (Security and Forensics), is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the test of my work.

This dissertation was prepared according to the regulations for postgraduate study of the Technological University Dublin and has not been submitted in whole or part for an award in any other Institute or University.

The work reported on in this dissertation conforms to the principles and requirements of the Institute's guidelines for ethics in research.

Signed: 

Date: 01 June 2019

ABSTRACT

Between January 2017 and March 2018, it is estimated that more than 1.9 billion personal and sensitive data records were compromised online. The average cost of a data breach in 2018 was reported to be in the region of US$3.62 million. These figures alone highlight the need for computer users to have a high level of information security awareness (ISA).

This research was conducted to establish the ISA of students in a university. There were three aspects to this piece of research. The first was to review and analyse the security habits of students in terms of their own personal device and examine their password habits, including their student account and their own personal accounts. The second was to assess and evaluate each student on a variety of scenarios related to security, using a quiz which had a series of multiple choice questions. Respondents were required to select the option that would be deemed the most secure. Finally, the third aspect of this research was to establish if respondents who had participated in ISA training in the past, scored higher in either the quiz or the assessment of their own device and password habits when compared with users who had not participated in any form of training. This was to determine if ISA training had any bearing on these types of behaviours.

The survey was opened up to students in TU Dublin (city centre campus) over a ten day period, with 752 participants taking part. The results of the survey were analysed using a number of statistical methods to identify if any significant differences existed between the various demographic groups when their own **security behaviours** and **knowledge of security best practices** were weighted and scored. Results from this research revealed that gender and student status were contributing factors to the scores obtained by students. The research also determined that ISA training also had a significant bearing on these two aspects.

Key words: *Information security awareness; Device habits; Password habits; security behaviours; security best practices; Demographic groups*

ACKNOWLEDGEMENTS

I would like to firstly thank my project supervisor, Dr Brian Keegan, for his guidance, contributions and support throughout last number of months completing this research.

I would also like to thank Dr Tom Clonan, for some kind words of wisdom and very useful advice in the early part of this research, along with Melda Slattery for her assistance with allowing the survey to reach a wider audience.

A big thank you to a number of colleagues who have helped along the way, including Richard Dunne, a colleague and fellow student, who battled through the same evening classes as myself.

Finally, I would like to thank my wife, Denise, and our three young children, for their patience, understanding and support throughout this research project.

Table of Contents

TABLE OF FIGURES

TABLE OF TABLES

1 INTRODUCTION

1.1 Background

The majority of university students use information technology for the purpose of completing their studies. A recent survey conducted by Educause found that over 86% of college students own a laptop and use it as their primary computer device for academic activities (Afreen, 2014). Most, if not all of these devices would be out of the control and administration of the college or university that these students are currently enrolled at.

As the rise of internet use and the use of online applications continues to grow, users who use these personal devices are becoming more vulnerable to security incidents. The overall range of threats users are being exposed to is growing at an alarming rate (Furnell, Bryant, & Phippen, 2007). Despite this increased use of technology in everyday life, users' behaviour with regard to data protection has not progressed at the same pace (Joinson, Reips, Buchanan, & Schofield, 2010)

A recent survey by PwC (Moran, 2018) shows that 61% of Irish organisations suffered cybercrime in the last two years, which is an increase from 44% in 2016. Research has demonstrated that students are particularly lax when it comes to the security related to their personal devices (Jones & Heinrichs, 2012). Although hardware and software security mechanisms are used by enterprise organisations and by end users to strengthen information systems against cyber-attacks, these systems can still be vulnerable to threats due to the user's risky behaviours. (Öğütçü, Testik, & Chouseinoglou, 2016).

With this increased use of personal devices in a university environment and with it, the increased exposure to threats, there is a need to evaluate the level of information security awareness of the student population. Students need to be aware of how to protect their information and systems from possible cyber-attacks or vulnerabilities.

1.2 Research Problem

Whitman and Mattord (2011) define Information security as "the protection of information and its crucial elements, including the systems and hardware that use, store and transmit that information". Students within a third level institution need to have a good understanding of these elements in order to prevent the loss of data and reduce the possibility of a security attack.

Colleges and universities in Ireland provide some level of training to students in relation to information security awareness. Despite implementing state-of-the-art technical controls, organisations still continue to experience security breaches (McCormac et al., 2017). One of the most important steps in developing these training programmes is to understand the level of information security awareness amongst the student population, in order to be able to customise the appropriate training programme. Ensuring students and staff within a third level institution are aware of these security best practices should reduce the amount of cyber related incidents across the board.

In order to implement a successful information security awareness campaign, it is essential to determine the security hygiene of all users beforehand. The purpose of this research will be to evaluate the level of awareness of information security of university students and to determine if there are significant differences in the information security awareness (ISA) levels between various demographic groups.

1.3 Research Question

The research aims to investigate and answer the following research question:

Are there certain demographic groups within a third level educational institute that have a lower level of information security awareness?

The main aim of this research will be to establish the level of security awareness between certain demographic groups within a third level educational institute. With the research question identified, several hypotheses were formulated which will be investigated during this research.

Hypothesis 1:

H0: When given a quiz relating to IT security awareness, there will be no significant difference in the mean scores for the various demographic groups

H1: When given a quiz relating to IT security awareness, there will be a significant difference in the mean scores for the various demographic groups

Hypothesis 2:

H0: When respondents' security behaviours and habits are weighted and scored, there will be no significant difference in the mean scores for the various demographic groups

H1: When respondents' security behaviours and habits are weighted and scored, there will be a significant difference in the mean scores for the various demographic groups

Hypothesis 3:

H0: There will be a significant relationship between users who have participated in information security awareness training and the scores they obtain when their behaviour and quiz scores are analysed

H1: There will be no significant relationship between users who have participated in information security awareness training and the scores they obtain when their behaviour and quiz scores are analysed

1.4 Research Objectives

The list of objectives for this research project are as follows:

Objective 1: Identify a list of areas and topics that should be used to assess the security awareness of third level students by reading additional research papers.

Objective 2: Assess the security awareness of students in a third level institute using an online survey, which will include a quiz.

Objective 3: Identify if certain demographic groups have a higher level of information security awareness than others

By completing these objectives, it should be possible to determine if certain demographic groups within a third level institute have a higher level of information security awareness than others. If these demographic groups can be identified, it may be possible for the university to target these specific groups with particular training programmes.

1.5 Research Methodologies

The research methodologies used in this study consisted of primary research and secondary research. The secondary research consisted of a comprehensive literature review which provided an insight into the existing background in the area of information security awareness. This included previous studies that have been undertaken in order to assess security awareness amongst end users as well as security best practices. It also focused on different types of surveys that could be used to acquire this information as well as a number of sampling methods that could be used. A number of statistical tools and methods were researched that could be used to prove or disprove the various hypothesis listed in this chapter.

The primary research consisted of using an online survey to collect demographic information from students within a third level institute, along with the behaviour and security habits of each student in relation to their personal device. The survey was

structured with a number of mandatory multiple-choice questions relating to their own device, which allowed respondents to answer these questions be selecting one of a number of pre-defined answers. Respondents were also asked to provide details relating to their password habits for their university student account as well as their own personal accounts.

After collecting details relating to their own behaviour and security habits, respondents were then presented with a quiz. A number of hypothetical scenarios were presented to each respondent, with each respondent asked to select the answer they deemed to be the most secure choice. A total of twelve questions were presented, with a score being assigned to each correct response. The questions used were formulated from existing literature and research carried out in this area.

1.6 Scope and Limitations

The literature reviewed for this research outlined a number of areas that should be assessed when evaluating the security awareness of users. This ranges from simple best practices when backing up data, keeping data secure, management and security of both laptop and mobile phones, to being able to spot phishing emails as well as awareness of social engineering attacks. It would be unachievable to assess every single aspect outlined in the literature, due to time constraints. This research focused primarily on the habits of the user in relation to the primary device they used for assignments, whether it be a PC or Mac laptop, as well as password hygiene, data protection, email best practices and awareness of phishing.

The survey only focused on certain demographic categories that were of interest to third level students, which included age, gender, student status, area of study and level of study. Areas such as income levels and area of employment were not included in these groupings due to the target population being students.

1.7 Document Outline

Chapter 2: Literature review

This chapter will review the existing literature in the area of information security awareness, which will include the risks associated with each area, including financial penalties that can be implemented under the General Data Protection Regulations (GDPR). It will look at previous studies that have evaluated particular demographics in relation to their security habits. It will review security best practices, including the recently changed guidelines for password policies. Best practices for surveys will also be examined, along with sampling methods.

Chapter 3: Design and methodology

This chapter will give an overview of the design and methodology of this study. It examines why a survey was used, reviews the various statistical tools and methods that were used to examine the research question.

Chapter 4: implementation and results

This chapter gives an overview of the results obtained from the survey. It starts by giving a breakdown of the various demographic groups and a summary of how each question was answered, using visual aids such as bar charts and pie charts to represent this data.

Chapter 5: Analysis and Evaluation

This chapter will discuss and analyse the various results obtained from the survey in order to determine if each of the null hypothesis outlined in the introduction chapter of this document can be either accepted or rejected. It will also give an overview of any significant findings that were identified as part of this research.

Chapter 6: Conclusions and future work

This chapter will summarise the research that was carried out. It will also outline the limitations of the research, discusses any possible future work and give some final thoughts.

2 LITERATURE REVIEW

2.1 Introduction

This chapter will define what information security awareness (ISA) is, along with a brief overview of why it is required in an educational institute. Studies have examined various types of information security awareness programmes in different environments, identifying what has been done to improve these types of programmes over the past number of years.

This chapter will also look at why ISA is important, outlining penalties that could be applied to organisations for failing to protect data or failing to disclose when a data breach occurred under GDPR. A number of studies were examined as part of this research to determine if there were differences in certain demographic groups when their level of ISA was assessed. This will include an analysis of previous security awareness surveys, phishing surveys and evaluating user security habits relating to device usage and password habits.

A number of research papers were looked at to examine what types of surveys could be implemented to best obtain this type of information, with a number of sampling techniques examined. A number of procedures and methods for determining the sample size for continuous and categorical variables were also examined.

2.2 What is Information Security Awareness (ISA)?

The concept of information security awareness is described in the literature to mean that users should be aware of security objectives (Siponen, 2001). In their research paper Hanus, Windsor, & Wu (2018) examined the multidimensional definition of security awareness. The most series deficiency in the literature they detected was the lack of consensus on what security awareness was. In the various papers where it was defined, the definitions were not consistent across the board. Albrechtsen (2007) describes it as an understanding of the importance of information security, along with the user related responsibilities. Others describe it as the level of knowledge and understanding of security issues within an organisation (Bulgurcu, Cavusoglu, & Benbasat, 2010) along

with an awareness of threats and security countermeasures and precautions (Ng, Kankanhalli, & Xu, 2009; Rhee, Ryu, & Kim, 2012). In each of the papers researched by Hanus et al., security awareness was not explicitly defined.

Information security awareness has also been described as ensuring that all employees in an organisation are aware of their role and responsibility towards securing the information they work with (Schultz 2004;Irvine & Chin 1998). It has also been described as a dynamic process, as the risks that users are exposed to continually change (Kruger & Kearney, 2006). Due to the ongoing change in technology, it is essential that any information security training that is offered to end users is continually measured, re-assessed and updated.

The increased use of technologies, along with the persistence of human weakness in information security continue to create new opportunities for cyber criminals. The threats related to human behaviour, such as social engineering, spear-phishing and cyber espionage have not changed over the past 20 years (Hanus et al., 2018). Security awareness amongst end users is often overlooked in an information security programme.

While organisations and enterprises around the world, including higher education institutes, expand their use of advanced security technologies as well as continually train their IT staff and security professionals, very little is done to increase the security awareness of their end users (Aloul, 2012). This in turn makes these end users the weakest link in the organisation. User behaviours are difficult to control, with the end user often being undertrained or unaware of what security is all about (Johnson, 2006). End user security awareness is a random variable that can be very difficult to characterise due to their individual nature (Dodge, Carver, & Ferguson, 2007)

While there is a risk to data theft involving personal data stored in social media accounts or access to financial data via online banking, educational institutes face other risks such as losing intellectual property or valuable research data, along with personal information relating to students, staff or faculty (Senthilkumar & Easwaramoorthy, 2017a).

A number of higher level educational institutes now recommend building security awareness training programmes for both students and staff, with the emphasis on the end user being kept up to date on all possible IT threats, allowing them to apply the security lessons in the most effective way (Piazza, 2006). Users can contribute and reduce these threats with several security actions taken on an ongoing basis. Some of these include locking their workstation when absent from it, password etiquette or password habits, cautious use of email and being able to identify suspicious emails, avoid using unlicensed software, keeping their operating system and software fully patched and up to date and reporting any information security breaches (Albrechtsen, 2007)

2.3 Information security awareness training programmes

The primary goal of a security awareness training programme is to make the end user aware of the various computer risks, how they can affect the organisation or educational institute the user is working for or enrolled in and to try and get the user to understand the importance of safe computer behaviour (Peltier, 2000).

A special publication on computer security training guidelines (Todd & Guitian, 1989, p. 8) completed by the National Institute of Standards and Technology (NIST) outlines the reason awareness training is required:

> *"Creates the sensitivity to the threats and vulnerabilities of computer systems and the recognition of the need to protect data, information, and the means of processing them"*

The presence of uneducated users in an educational institute, whether they be staff or students, makes them a prime target for hackers. Aloul (2012) outlines recommendations in his research paper that security awareness should be done on a regular basis, but more so, that the method for preparing the awareness training is very important in most cases. The content needs to be customised for different users, but it should cover the organisations IT security policy. The other major factor is how the awareness material is delivered to the end user. One of the key findings in this journal article is that enterprises should adapt a proactive rather than a reactive approach to security awareness (Aloul, 2012).

2.4 Why information security awareness is important

A recent survey conducted by EY[1] found that there were over 1.9 billion personal records breached between January 2017 and March 2018 (van Kessel, 2018). This survey also identified that more than 87% of organisations surveyed did not have the sufficient budget in place to provide the levels of cybersecurity and resilience they wanted.

A number of different threats have been identified in the literature which have been described as the most common ways cybercriminals will try to steal data or gain unauthorised access to a system. It has been well established that the weakest link in any organisation is the end user in terms of computer security countermeasures (Rhodes, 2001). Social engineering is a common method used by attackers which involves persuading individuals that the perpetrator is someone other than who he/she really is (Mitnick & Simon, 2011). These social engineering attacks can involve taking advantage of a known vulnerability in a system, or by carrying out a phishing attack on a user.

In recent years, a more dangerous type of cyber-threat has emerged which is known as ransomware. Ransomware is a type of self-propagating malware which uses encryption to hold a victim's data ransom until a payment is made, usually in the form of a crypto-currency (Chen & Bridges, 2017). One of the most well-known and much publicised cases of ransomware was "WannaCry", which was a large scale cyber-attack that occurred in May of 2017 which targeted Microsoft Windows systems, infecting more than 230,000 computers in over 150 countries (Ehrenfeld, 2017). Although a number of sectors were affected by this attack, the National Health Service in Britain were significantly impacted, with more than 60 NHS trusts hit with this attack. This prevented many facilities from accessing patient records, which led to significant delays and cancellations of non-urgent surgeries and patient appointments (Collier, 2017) . Ehrenfeld (2017) outlined that the entire situation was preventable, as Microsoft had released a critical patch in March of 2017, which once applied, removed the vulnerability required for this malware to propagate from machine to machine.

[1] https://www.ey.com/en_gl/advisory/global-information-security-survey-2018-2019

2.4.1 Financial penalties

With the introduction of the new General Data Protection Regulations (GPPR) in May 2018, authoritative entities now have greater legislative powers to fine organisations in the event of a data breach (Albrecht, 2016). GDPR's primary objective is to strengthen and harmonize data protection for individuals as well as to simplify regulatory environments for organizations (Zerlang, 2017)

Companies and organisations now have to notify EU authorities of a data breach within 72 hours (Sharf, 2016). Companies have a responsibility to ensure personal information is kept safe. Penalties for breaching the GDPR comprise of up to 4% of the previous year's profits (McCall, 2018). For minor breaches, organisations can be fined up to 2% of their worldwide turnover (Tankard, 2016)

Zerlang (2017) found that in the past, organisations have chosen to secure only the most mission critical elements of their business. In today's digital landscape, there now exists a greater number of threat actors, methodologies and entry points. This means that any networked device an employee or a student uses within the organisation now represents a potential threat. With this increased financial penalty associated with possible data breaches, it is now more essential that employees and students within a third level institute are aware of the various security risks associated.

2.5 Previous Research

Prior research has been carried out to evaluate and compare security habits of users. Lon, Reeder and Consolvo (2015) compare the results of two separate online surveys, one with 231 security experts and the other with 294 non-security-expert internet users. The results show that there were discrepancies in the behaviour of both groups in relation to security practices. The results of the survey showed that 73% of security experts used a password manager on some of their accounts, compared to just 24% of non-security experts. An assumption of the low adaption rate of password managers by non-security experts was possibly due to the lack of understanding of its security benefits.

In relation to two-factor authentication, the results are similar to the uptake of using a password manager. 89% of security experts claimed that they used MFA on at least one of their online accounts, compared to 62% of non-security experts. It was highlighted by experts in the survey that the majority of non-expert users do not understand two-factor very well, with some needing further instructions on how it works (Ion et al., 2015). This research paper also looked at aspects relating to safe web browsing, particularly if users checked if the web-page they were accessing used HTTPS or not. The results showed that 82% of experts said they check this often, compared to 36% of non-expert users.

The overall results showed that the security experts surveyed regarded installing updates on systems, using a password manager, using strong passwords and two-factor authentication as the top pieces of advice to give to non-tech-savvy users, whereas non experts considered that installing anti-virus, frequently changing passwords and only visiting trusted websites were considered effective measures

Aytes and Connolly (2004) conducted a survey of 167 undergraduate students at two large public universities to determine the frequency in which they engaged in five common but unsafe computing practices, including sharing passwords, opening unknown attachments and not backing up data on a regular basis. The results of the survey outlined that 22% had reported that they did not share passwords, with over 51% claiming they had never or rarely changed their password after creating an online account. Additionally, only 38% of respondents claimed to back up their data "frequently" or "all of the time". Their findings suggested that users will continue to use risky behaviours, regardless of the risks being outlined to the user. Their findings also suggest that it is unlikely that computer users will change their behaviour in response to simply being provided with additional information relating to security risks and best practices.

Rezgui & Marks (2008) carried out a study to explore factors that affect information security awareness of faculty staff in a university, which also included information systems decision makers. Their case study revealed that factors such as consciousness, social conditions and cultural assumptions and beliefs affect university staff behaviour.

They also found that only a small percentage of the universities provide some form of security awareness training. This statistic is confirmed by a quantitative survey carried out on over 400 higher education institutes conducted by Luker & Petersen (2003).

The use of a video game for cyber security training and awareness was put forward by Cone, Irvine, Thompson, & Nguyen (2007). They argue that many forms of training fail because they are rote and do not require the user to think about and apply security concepts. They proposed using a flexible highly interactive video game as a security awareness tool. In comparison to other games that have been developed, Cone et al. argue that no other developed games combine the human and technical factors associated with an IT environment. The results indicated that the game is being successfully utilized for information assurance education and training by a variety of organisations, although the paper did not have a way of evaluating if the format used for training users was more effective than standard practices.

McCormac et al. (2017) discussed the relationship between individuals' information security awareness (ISA) and individual variables, such as gender, age, personality and risk taking propensity. They carried out a survey of over 500 working Australians. The results obtained showed that older adults had a higher ISA score compared to younger adults, and in terms of gender, females scored slightly better than males. These gender differences matched the results of a similar study carried out Sheng et al. (2010).

Another approach to determining the information security awareness of users was carried out by Thomson & von Solms (1998) which found that in order to be more effective, the ISA should be tailored to address specific groupings of employees. In their research paper, Kim (2014) outlined a series of recommendations for developing security awareness training for college students. This paper outlined that there were two possible approaches to improve information security in an organisation. The first being a "sanctions-based approach" where fear of possible sanctions would determine whether the end user would comply with such policies. Due to the surge in students using their own devices in a third level institute, this approach would not be applicable to this research study. The second approach is to persuade end users to make the right choices through Information security awareness training (ISAT).

Wilson & Hash (2003) identify a number of topics that could be used in a security awareness campaign. A number of these will be used as a basis to determine the topics that staff and students will be asked in the online survey, as well as the areas that will be used in a series of interviews with IT security experts.

A number of surveys have been carried out in various areas that assess the security awareness of individual users. Some of the surveys carried out evaluated awareness and behaviours of a number of different areas, whereas some surveys just focused on particular aspects that make up the overall understanding of security awareness, such as susceptibility to phishing, or whether or not users regarded installing OS updates as an important aspect. Each survey focuses on certain aspects, which will now be looked at in more detail

2.5.1 Security Awareness surveys

A study carried out by Albrechtsen & Hovden (2010) demonstrated that local employee participation, collective reflection and group processes produce changes in short term security awareness and behaviour. In this study, a survey was issued to all users one month before an intervention took place. This survey consisted of a series of questions relating to different statements on information security topics, which the respondents were asked to agree or disagree with based on a 5-point Likert scale.

Participants were divided into three groups, where two of these groups were invited along to an intervention, with the third group being set as the control group. This control group would receive both surveys, but not be involved in the group discussion. The intervention was structured as a forum of discussion, with the participants encouraged to contribute with their thoughts about information security. A number of animated videos were shown throughout the intervention, which covered a number of scenarios followed by a group discussion.

A second survey was sent to respondents who participated in both the intervention and those in the control group a month after the intervention took place, with a third and final survey sent to participant a year after. This was to evaluate the stability of the awareness training. The results observed showed that within the 2 groups that were involved with

the intervention, 66% (group 1) and 67% (group 2) reported that their awareness had changed within the past year, compared with 27% in the control group.

This study demonstrated that locally based employee participation, collective reflection along with group interaction create changes in information security awareness and behaviour at an individual level. Qualitative data gathered from the employees after this study was completed identified that the success of these workshops was down to the way the information was presented in a relaxed and humorous atmosphere.

2.5.2 Phishing surveys

Sheng et al., (2010) carried out a roleplay survey on over a thousand students in a university which was used to study the relationship between demographics and phishing susceptibility and the effectiveness of several anti-phishing educational material. In this online study, participants completed a role-play task where they were shown emails and websites which may or may not be phishing attempts. Participants were then given one of several forms of training, before then been given a second role-play task to once again to assess their behavioural susceptibility to phishing.

Their results showed that women were more susceptible to men and users in the age category of 18-25 were more susceptible to any other age group. Although it was established that the use of educational materials to help users identify phishing sites reduced users' tendency to enter details on these sites, it did decrease the participant tendency to enter information on legitimate websites. Overall, prior to being shown the training material, participants on average fell for 47% phishing websites, whereas after the training was provided, this number reduced down to 28%. These figures are comparable with the results obtained by a similar study involving another role-play survey by Kumaraguru, Sheng, Acquisti, Cranor, & Hong (2010).

2.6 Analysis of security habits and behaviours

2.6.1 OS and Software Updates

Vaniea, Rader, & Wash (2014) identified three reasons as to why end users failed to install system updates, or just didn't bother with the process: (1) Participants found that security updates often bundled with other undesirable features. (2) Users also had difficulty in assessing the value of an update on the system and (3) some users were confused as to why updates were needed at all. As discussed earlier in this chapter, in their research paper, Ion et al., (2015) found that 35% of IT experts surveyed identified the importance of installing OS updates, whereas only 2% of non-experts mentioned this when surveyed.

2.6.2 Anti-virus and Anti-Malware

With the advances in security technologies, a lot of computing behaviours such as patch management and anti-virus updates are now being automated to reduce the expertise required by the end user as well as the time burden (Herath & Rao, 2009).

A survey carried out on university students by Katz (2005) found that only 27% of students surveyed agreed or strongly agreed with the statement "using the anti-virus program loaded on my PC, I always execute an anti-virus scan of my computer at least once a week". Another study carried out by Senthilkumar & Easwaramoorthy (2017b) surveyed a number of third level students to establish their behaviours when it came to Anti-virus software. They found that although over 70% of the students were aware of basic virus attacks and had anti-virus software installed on their personal devices with 11% of these admitting that they did not update their antivirus software or did not know how to.

A proof-of-concept field study was carried out by Lalonde Levesque, Nsiempba, Fernandez, Chiasson, & Somayaji (2013) to examine interactions between users, anti-virus/anti-malware software and malware as they occur on deployed systems. This four-month study involved providing laptops to 50 subjects which were all setup with the same configuration and software to monitor for malware infections. During this study, 380 files were detected on 19 different user machines by the pre-installed anti-virus

software, indicating that 38% of the test population were exposed to malware. These results would indicate that if they were representative of the entire user population, almost 1 out of every 2 newly installed machines would be infected within the first 4 months of their use if anti-virus software was not installed on the device.

Kharraz, Robertson, Balzarotti, Bilge, & Kirda (2015) discuss the results of a long-term study on ransomware attacks that have been observed between 2006 and 2014. This class of malware, also known as scareware, locks the user out of their data until a ransom has been paid. One type of ransomware that was analysed in this study was the crypotlocker ransomware, which managed to infect over 250,000 computers around the world. Analysis was carried out on 1,872 bitcoin transactions that were used during the crypotlocker attack, which shows that new bitcoin addresses were used for each infection to keep the balances of each bitcoin address low. This indicated that cybercriminals were starting to use new evasive techniques to better conceal their criminal activity (Kharraz et al., 2015)

Another cause for concern is the recent trend in the use of fake Antivirus software being advertised on bogus sites. Hackers are using new and ingenious methods in order to gain access to other users' systems. Over the past few years, a number of bogus websites offering free anti-virus software have been identified which can end up infecting an end users' computer, resulting in their personal data being compromised (Safa, Solms, & Futcher, 2016).

2.6.3 Password hygiene and password habits

Over the past number of years, the number of passwords that users have to create and remember has risen considerably, with users accumulating more and more accounts and services. As a result, users are now required to remember multiple passwords which can introduce risky password behaviours (Woods & Siponen, 2019). This can include password reuse, writing passwords down, choosing weaker passwords that are easier to remember and not changing passwords regularly (Guo, 2013). One recommendation to overcome these risky behaviours is to use a password manager. Although password managers have been around since the early 90's, the uptake with users have been limited, with some users believing they are vulnerable to attacks (Woods & Siponen, 2019). Das,

Bonneau, Caesar, Borisov, & Wang (2014) conducted a survey to understand users' behaviours when they were creating passwords across multiple sites. The results showed that out of 224 participants, only 6% of them chose to use a password manager.

Stobert & Biddle (2014) examined user behaviour of managing passwords. This study involved a series of interviews with 27 university students to determine how users coped with having to deal with a large number of passwords. They found that all but one user interviewed re-used the same password on multiple sites. Most of the participants appeared unaware of prominent password managers, with some participants expressing distrust in this type of software. Another finding of this study was that most of the users had little understanding or knowledge of using single sign-in where it was provided, which would address the issue of having to create and remember new passwords.

Wash, Rader, Berman, & Wellmer (2016) examined a series of self-report survey responses with some 134 participants to determine how frequently entered passwords are re-used across multiple sites. As well as the survey, users installed custom written log data collection software on their personal computers so a comparison could be done on the user's self-reported beliefs and behaviours with their actual password characteristic and re-use. This research determined that users tend to re-use passwords that they have to enter frequently, and those passwords tend to be among the users' strongest passwords. More interestingly, because the software was able to log user's password entries, they could also see where a user had entered an incorrect password on a different site, in the most cases the user would use their "go-to" password to try and authenticate on that site.

A number of other studies carried out determined that users have a similar number of distinct passwords. A large scale study of password habits of more than half a million internet users by Florencio & Herley (2007) examined the password use and re-use habits . Users opted in to installing client software that would scan HTML pages for submitted passwords for each URL they accessed. If the software found a password entry, it would hash the password and store it in the protected password list (PPL) within Microsoft Windows. The software also recorded the bit-strength of the password. From this, it was possible to determine which passwords contained (1) lowercase only, (2)

lowercase and digits, (3) lowercase, uppercase and digits and (4) all four types. Unless a particular website forced the use of these types of characters, their data showed that the majority of passwords used contained only lowercase letters. When users were forced to use stronger passwords, there was a tendency to only use longer lower-case passwords, and not use any of the other character types.

Within the same study, Florencio & Herley (2007) were also able to determine the number of times passwords re re-used across multiple sites. Over the course of the two-month study, they determined that users re-used the same password at just under 6 distinct login sites. It was also found that users averaged 6.5 distinct passwords.

2.7 Password guidelines

The National institute of standards and technology (NIST) released a publication in 2017 called the NIST Special Publication 800-63B, outlining updated recommendations for password length and complexity requirements (Grassi et al., 2017). In terms of complexity, password composition rules are commonly used in order to increase the difficulty of guessing a user-chosen password. This research found that analyses of breached password databases revealed that when complexity was enforced as a requirement for user-chosen passwords, the user setting the password responded in very predictable ways to the requirements imposed by these rules. For example, if a user who chose “september” as their password would be likely to choose “September1” if they were required to include an uppercase and number. Similarly, if a symbol was required, they would likely choose “September1!”

Due to these findings, Grassi et al. (2017) found password length to be a primary factor in characterizing password strength. Users should be encouraged to make passwords as lengthy as they wish, within reason, as long passwords could conceivably require excessive processing time to hash (Grassi et al., 2017).

2.8 Surveys

Andrews, Nonnecke, & Preece (2007) identified five methodological components that were critical to successful web-based surveys. These include (1) survey design, (2)

subject privacy and confidentiality, (3) sampling and subject selection, (4) distribution and response managements and finally (5) survey piloting.

Over the past 20 years, the use of the internet has become a lot more widespread, with many social scientists conducting surveys through this medium compared to face-to-face surveys or telephone surveys (Fraley, 2004). Online surveys have the potential to reach a much larger, more diverse population and may be as effective as standard mail surveys (Gosling, Vazire, Srivastava, & John, 2004). They also have the potential to achieve sample sizes that exceed mail or telephone surveys. Online surveys are probably the most cost effective means of data collections when the target population is students within a college campus (Matsuo, Mcintyre, Tomazic, & Katz, 2004).

2.8.1 Why use an online survey

Traditional survey literature identifies three possible response behaviours; Unit-non response, Item non-response and complete response (Bosnjak & Tuten, 2001). One advantage of using a web-based survey is the ability to capture data about a respondent's answering process.

When designing a survey, the order of topics can have a significant impact on the dropout rate. Frick, Bächtiger, & Reips, (1999) investigated the effects of asking for personal information at the beginning of a survey compared to it being asked at the end of a survey. Surprisingly, drop-outs were significantly higher when this information was asked at the end of the survey (17.3% compared to 10.3%).

Dillman (2011) discusses the importance of not alienating users who are uncomfortable with using the web. It was identified that the use of pull-down menus, unclear instructions, along with a lack of navigation aids may result in novice web users from completing a survey.

Another part of this research examined the use of incentives on response. (Frick et al., 1999) concluded that when the chance to win a prize was offered as an incentive to complete a survey, this resulted in a lower drop-out rate compared to when no prize was offered. The opposite was found by Tuten, Bosnjak, & Bandilla (1999). They found

that the number of non-responders was considerably higher when a chance to win a prize was offered than in cases where the user was advised that their participation in the survey was contributing to scientific research. For this purpose, the chance to win a prize for completion of the survey was not offered to participants.

2.9 Sampling Techniques

Sampling is related to the selection of a subset of individuals from within a population in order to estimate the characteristics of the whole population (A. S. Singh & Masuku, 2014). It can be difficult to study the entire population as it can be costly, time consuming and complex (S. K. Singh, 2015). There are two major categories of sampling methods that exist; probability sampling and non-probability sampling. These categories contain a number of sampling techniques, which are listed in the table below.

Probability sampling	Non-probability sampling
Simple random sampling	Convenience sampling
Systematic random sampling	Judgment sampling
Clustered Sampling	Snow-Ball
Stratified Sampling	

Table 2-1: Sampling techniques

2.9.1 Probability Sampling

Probability sampling is where all subjects in the target population have an equal chance of being included in the sample (Elfil & Negida, 2017). Samples which are selected using these methods are more representative of the target population. One of the main disadvantages of using probability sampling techniques is that it can be tedious and time consuming, especially when the population size can be quite large. Simple random sampling is the most common type of probability sampling. This method is used when the whole population is accessible. From this population, each member is assigned a number and a lottery method is used to determine which subjects are included in the random sample (Elfil & Negida, 2017). Systematic random sampling is similar to simple random sampling, where the first unit of the sample is selected at random, but subsequent subjects are selected based on a systematic rule, using a fixed interval (A. S. Singh & Masuku, 2014).

Clustered random sampling, also known as Multistage sampling, is generally used when the population size is extremely large. Using this method, the population is divided into different by geographic location into clusters. A full list of clusters is then put together, with investigators using a lottery method to select which clusters will be used in the sample. Once this is decided, a full list of individuals within these clusters is listed and another turn of random selection is made on these individuals to generate a sample size (Elfil & Negida, 2017). Finally, stratified random sampling can be used if the population is heterogeneous. Using this method, the entire heterogeneous population is divided into a number of homogenous groups. These groups are generally referred to as Strata. Each of these groups is homogenous within itself. Units are then sampled at random from each of these stratums (A. S. Singh & Masuku, 2014)

2.9.2 Non-probability Sampling

Non-probability sampling is when the sampling population is selected in a non-systematic process, which does not guarantee an equal chance for each member of the target population to be included in the sample. Convenience sampling is also known as haphazard sampling or accidental sampling. This sampling technique is the most widely used method in clinical research (Elfil & Negida, 2017). Using this method, subjects are selected based on their geographical proximity, availability at a given time or the willingness to participate (Dörnyei & Griffee, 2010) meaning this method is quick, convenient in inexpensive (Elfil & Negida, 2017). The main assumption associated with convenience sampling is that the members of the target population are homogeneous and there should be no significant difference in the research results compared to that of a random sample (Etikan, Musa, & Alkassim, 2016).

Judgement sampling, which is also known as the purposive sampling technique, is the deliberate choice of a participant, due to the qualities the participant possesses. The researcher will assume specific characteristics for the sample (e.g. a male/female ratio of 3/1) which will allow them to judge the sample to be suitable for representing the population (Elfil & Negida, 2017). Teddlie & Yu (2007) identified that this method has been widely criticized due to the likelihood of bias by investigator judgement.

Another method used can be snow-ball sampling. Using this method, the investigator asks each subject to give them access to one of their colleagues from the same population. This method is generally used when it is difficult to locate the population in one place, or if the population is hard to reach (Elfil & Negida, 2017)

2.10 Sample Sizes

There are a number of procedures and methods for determining the sample size for continuous and categorical variables. Bartlett & Ik (2001) described the procedures originally outlined by Cochran (1977) and focus on the areas that need to be taken into consideration when calculating the sample size. It was outlined that Cochran's (1977) formula uses two key factors; the risk a researcher is willing to accept, which is known as the margin of error and the probability that differences revealed by these statistical methods really do not exist, which is known as the alpha level.

In most education research studies, the alpha level used in determining sample sizes is either 0.5 or 0.1 (Ary, Jacobs, & Razavieh, 1996). When using Cochran's formula, Bartlett & Ik (2001) outline that the alpha value is incorporated into the formula by utilizing the t-value for the alpha level selected. For a confidence level of 95%, the t-value is equal to 1.96, whereas for a confidence level of 99%, the t-value is equal to 2.58. The second item to consider when using Cochran's formula is the margin of error. For categorical data, a 5% margin of error is acceptable (Krejcie & Morgan, 1970). Bartlett & Ik (2001) also describe that when using Cochran's formula, if the figure obtained exceeds 5% of the population, Cochran's (1977) correction formula should then be used to calculate the final sample size.

2.11 Gaps in the research

Previous security awareness research has examined the individuals' information security awareness and individual variables (McCormac et al., 2017), but only focused primarily on the users gender and age. Within this study, the users were not asked if they had partaken in security awareness training beforehand but were simply assessed on the level of security awareness they portrayed through means of a survey. This survey was aimed at working Australians.

Drevin, Kruger, & Steyn (2007) used a value focused approach in their paper to identify key areas of ICT security. This was done using a series of interviews with various stakeholders, which although aided the University in providing a sustainable ICT security service to all staff and students, it did not determine which staff or students required a customised/focused set of security awareness training.

While a number of studies have compared the habits of security experts with non-security experts, with some examining certain characteristics of password usage and habits, it remains unclear if any one type of demographic has a higher security awareness than others. In the majority of these studies, the sample size surveyed has been considerably low. However, some of the studies indicated that factors such as gender and education levels may have a significant difference, which merits further investigation. Little research has been done to assess both the knowledge of security awareness and the behaviours of students in higher education.

2.12 Summary

In this chapter, a variety of literature relating to information security awareness was examined. A number of definitions of information security awareness were outlined, along with why ISA is important and what financial penalties exist when a company or organisation suffers a possible data breach.

A number of studies were looked at to determine why ISA programmes sometimes fail and what improvement have been recommended by experts in this field. Furthermore, the security habits of a number of demographic groups were compared between a number of previous studies, particularly in the area of device security and password hygiene. The take-up and use of password managers and multi-factor authentication were also examined.

This chapter also examined the advantages of using an online survey, various sampling techniques that can be used and finally research was carried out on how to determine the appropriate sample size required from the student population.

3 DESIGN AND METHODOLOGY

3.1 Introduction

This chapter will discuss the design and methodology used in order to evaluate the security awareness of university students within TU Dublin, city centre campus. A breakdown of the various demographic groups is described in detail, with the reasons behind selecting each categorical data to represent this data explained. An overview of the survey design and why a survey was used to gather this information is discussed. Details of the pilot study are outlined, along with the sample size formula used to determine the appropriate sample size that was required to give a 99% confidence level. Statistical tools & methods are explained, with a table outlining the scoring conversion used to assess the behaviour of respondents in relation to their security habits.

3.2 Design Overview

Little research has been carried out to assess the various demographic factors and how they differ in relation to information security awareness (ISA). The survey collects demographic information, gather details about each respondent's security habits (device usage, password habits) and then assesses their awareness using a quiz. Although the quiz will determine if the respondents are aware of best practices in the area of ISA, the assessment on their existing habits will underpin this to determine if they actually implement these best practices.

TU Dublin is newly created university, which was formed on the 1st of January 2019 when three existing Institutes of Technology based in Dublin were merged. It previously consisted of Dublin Institute of Technology, Institute of Technology Blanchardstown and Institute of Technology Tallaght (TU Dublin, 2019). These three campuses are now formally identified with the campus names outlined in Table 3-1 below.

Original name	New Name
Dublin Institute of Technology	TU Dublin – City Centre Campus
Institute of Technology Blanchardstown	TU Dublin – Blanchardstown Campus
Institute of Technology Tallaght	TU Dublin – Tallaght Campus

Table 3-1: Original and new names for TU Dublin

This research will focus only on students based in TU Dublin City Centre Campus, which has a population of approximately 19,528 students. These figures were provided by the Strategic Development office based in TU Dublin City centre campus. Any future reference to TU Dublin in this document relates only to the TU Dublin – city centre campus, unless otherwise stated.

3.2.1 Demographic Overview

The demographic data collected in this survey was structured in a way that it can be regarded as categorical variables. Respondents are categorized based on answered given through the survey. Each respondent can be assigned to one category (e.g. full-time or a part-time student) but cannot be part of more than one category per demographic group. In order to capture the demographic information of each respondent, a number of categories were defined for each question, with each respondent able to select only one category per question. In terms of gender, respondents could select if they were "male", "female", "rather not say" or "other". If respondents selected the option for "other", they could then type in whatever gender they wish to be identified by.

For the category of Age, the demographic set was divided into the following categories:

1. 17-19
2. 20-21
3. 22-23
4. 24-27
5. 28-34
6. 35-44
7. 45-54
8. 55+

A breakdown of the student ages was also provided by the Strategic Development office within TU Dublin. These figures can be seen in Appendix B of this document. Due to the high number of students in the age range of 17-24, the decision was made to narrow these groups into two-year intervals. The number of students over the age of 35 gradually declines to less than 200 per increment. Due to this decline, the age ranges higher than 35 were placed into 10-year intervals.

The city centre campus is primarily made up of four colleges, which are listed below:

1. College of applied arts and tourism
2. College of business
3. College of engineering and built environment
4. College of science and health

There are also a number of small schools, such as Learning and teaching technology centre (LTTC) and a graduate school. Students could select one of the above listed four colleges are their primary area of study or could opt to manually enter in the area of student they were involved with. A number of respondents entered in the course code or specific area of study, such as photography, when completing this question. These manually entered details were re-classified into the appropriate college once the survey had been closed. Students could select if they were a full-time or part-time student, and also declare at what level of study they were currently at from the following list:

1. 1st year undergraduate
2. Year 2, 3 or 4 undergraduates
3. Graduate (Masters)
4. Post graduate (PhD)
5. Apprentice / Trades

3.3 Student device assessment

Although the helpdesk within TU Dublin does not troubleshoot or repair student owned devices, students generally contact the helpdesk for advice and assistance with using services provided by the University, such as Wi-Fi, Student printing and obtaining

access to cloud services, such as Microsoft Office 365 products, which students are permitted to install on their own devices. Over the course of a 2-week period, the helpdesk took record of the various devices that students required assistance with. A high proportion of these included students enquiring about connecting their mobile phone to the wireless service in the college.

In relation to personal devices that the students would use to complete University assignments (i.e. laptop devices or tablet devices), a total of 23 students called into the helpdesk in the first week and a further 27 called to the helpdesk in the second week. The following device types that the students were looking for assistance with were identified by the helpdesk.

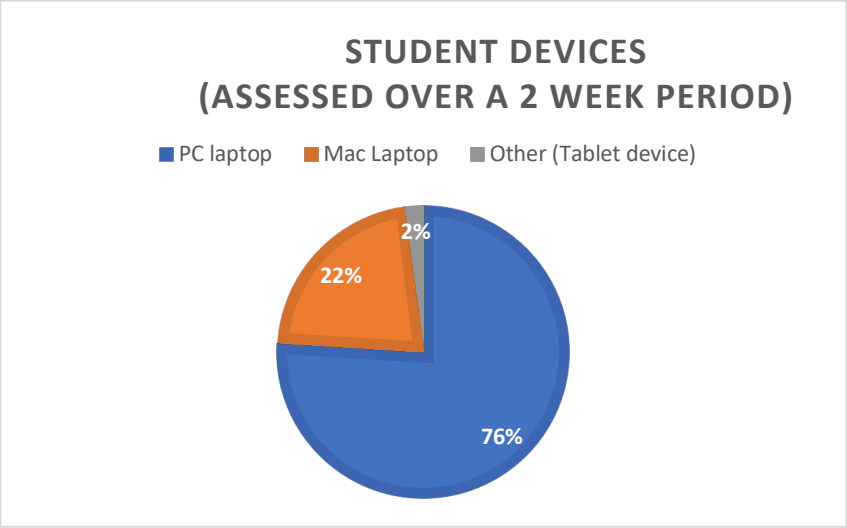


Figure 3.1: Student device usage

A member of the student helpdesk advised that these figured were consistent with what would be generally used by students throughout the campus. Based on these findings, it was appropriate to assess habits by users on these two types of devices within the survey.

3.4 Survey Design and Responses

3.4.1 Why a survey was used

In order to ascertain the security habits and the IT security awareness of the students within TU Dublin, we need to gather a variety of data from the student population. One of the quickest ways to gather this type of data is to use an online survey. This allows the researcher to make an inference about the wider population, which is known as the population of interest (Kelley, Clark, Brown, & Sitzia, 2003)

The main advantages of using an online survey over mail surveys are that there is no need for printing or postage, which can be a huge cost savings. Other advantages are the speed at which data can be collected is significantly faster than mail surveys and the precision of data compilation. There are also some disadvantages of using an online survey (Matsuo et al., 2004). These can be lower response rates, non-responses as well as the non-representativeness of the sample population, which can result in a lack of validity of the data collected.

In order to try and obtain a high response rate to the survey, it was important to keep the survey short and not make it too burdensome on the users partaking in it. Research carried out by Galesic (2006) examined the effects of interest and burden experienced by users who participated in an online survey. It was determined that incentives, short announced length or general interest in the topic were all influential in the user's preference to complete the survey.

3.5 Overview of survey

The survey was divided into seven sections. The first section recorded the various demographic information of each respondent. Section two to five recorded information relating to the respondent's behaviours, which were broken down into device usage habits, password habits, understanding of data protection and understanding of wireless technologies. Figure 3.2 below gives an overview of each of these sections.

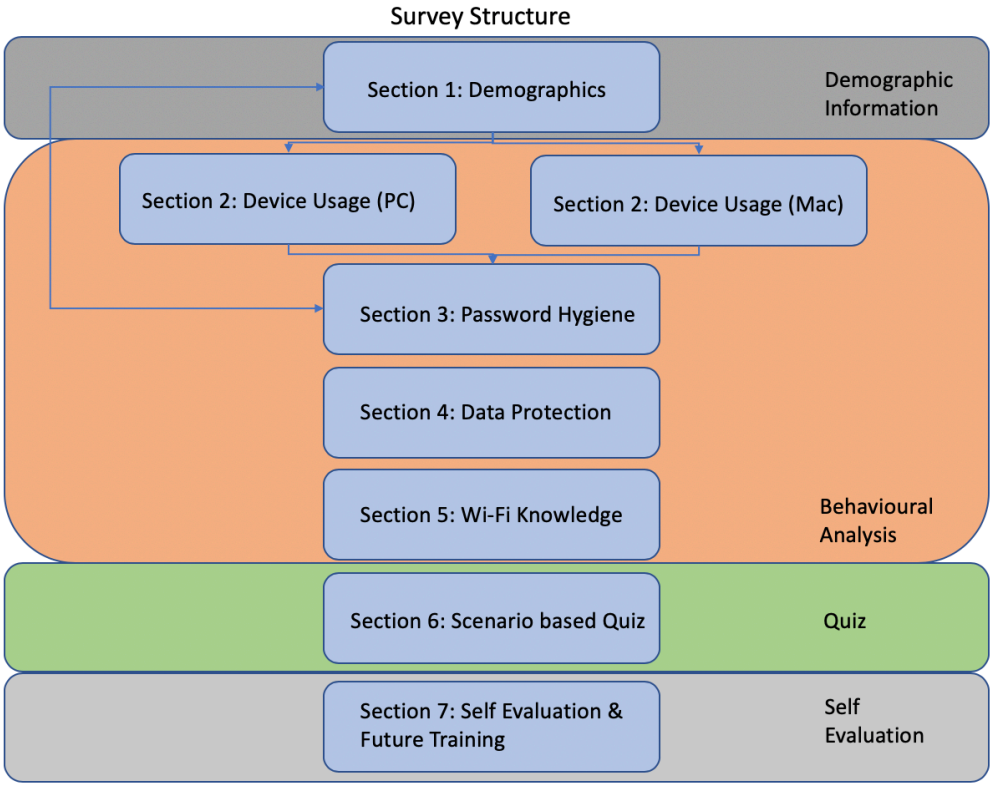


Figure 3.2: Survey Structure

Section 1: Demographics

Section 1 of the survey collected demographic information relating to the following key areas:

- Gender
- Age range
- Current level of study
- Part-time / full-time study
- Education discipline

It also established if the participant has undertaken any security awareness training provided by TU Dublin or elsewhere within the past 2 years or if the participant was aware that security awareness training was available for users to avail of. Participants were then asked to rate their IT competency levels on a scale from 1-7, as well as rate their IT Security awareness on a similar scale. Respondents were also asked if they had ever experienced a security breach. This information could be used to determine if users previously involved in a security breach would score higher due to the fact that they have previously been targeted by cyber criminals.

The last question respondents were asked in this section was related to what type of personal device the user owned and used as their primary device for completing college assignments. The user was presented with a choice of either a PC laptop, an Apple Mac laptop, something other than a PC or an Apple laptop or that they did not own a device. If the respondent stated they owned and used a Windows PC laptop, they would be presented questions related to Windows devices. Likewise, if the respondent stated they owned and used an Apple Mac laptop, they would be presented with questions related to a Mac laptop.

If respondents have stated they used something else other than a Windows PC or an Apple Laptop or that they did not own a device, they would skip to section four, which was related to Password Hygiene.

Section 2: Device Usage

Section 2 of the survey collected information relating to the type of device the user used as their primary device for completing college assignments. These questions determined the following:

1. The OS version running on the device
2. If the device was password protected
3. If the device was encrypted
4. If the device had antivirus installed
5. How often the user installed OS/Security updates on the device
6. How often the user updated software on the device
7. If the device had a firewall enabled
8. If the primary account on the device was an admin account
9. If the user allowed other users to use their device
10. If the user regularly backed up the data on their device

If the user stated that they had Anti-virus installed on the device, they were presented with an additional set of questions, which asked the following:

1. How often the user updated Anti-virus definitions on the device

2. If the user regularly scanned their device for viruses

Section 3: Password Hygiene

This section of the survey was used to determine the password habits of each respondent. Each student within TU Dublin is given a username and password when they enrol as a student. The username is the student number, with the password being a randomly assigned password. Although students are encouraged to change their password when they register, students are not forced to change it upon login. Due to this policy, students could complete a full 4-year course without having to change their password once.

This section questioned the students' behaviour regarding their TU Dublin account and also assessed their habits in relation to their own personal online accounts, such as social media accounts or additional emails accounts. The following questions were presented to the respondents in the survey:

1. How often they changed the password on their student account
2. How long the password was for this account
3. How complex this password was?
4. How often they changed their password for other accounts they used
5. If they used the same passwords on multiple sites
6. If they regarded their password as strong
7. If they used a password manager to store their online account passwords
8. If they allowed their web browser to store their passwords
9. If they were aware of what MFA was (Multi-factor authentication) and if they used it

Previous studies by Stobert & Biddle (2014) examined the password length and how often users changed their passwords. The results of this survey could be used as a comparison to determine if this type of behaviour was consistent with previous results

Section 4: Data storage

In this section, users were asked a simple question of whether or not they used a USB key or an external hard drive to store data. If the respondents stated that they did, they were then presented with a second question asking if this device was encrypted.

Section 5: Wi-Fi Knowledge

This section presented a set of questions relating to wireless network connectivity. Respondents were given the following set of questions:

1. Has the student ever connected to an open/insecure wireless connection?
2. Has the student ever checked their online banking or sent email over this type of connection?
3. If they were aware, using appropriate tools, that a hacker could intercept their wireless traffic over an insecure/open connection

Section 6: Quiz

This section of the survey consisted of a number of multiple-choice questions, where the user would be awarded 1 point for each correct answer, with a total of 12 points that could be achieved by each user. These multiple-choice questions placed the respondent in a particular scenario and presented them with a series of possible answers. Research has indicated that when using surveys, respondents may tend to select the first few response options when given a multiple choice question. (Choi & Pak, 2004). This phenomenon is known as primacy bias. To eliminate this type of bias, multiple-choice answers were set to be displayed in a different order for each respondent that participated in the survey. The questions in this section covered aspects related to the following:

1. Phishing attempts & email (3 questions)
2. Wireless technology (1 question)
3. Passwords/MFA (4 question)
4. Data Protection (4 question)

A complete list of these questions can be found in Appendix A of this document

Section 7: Self-evaluation and future training

After each respondent completed all of the multiple-choice questions, they were then asked to assess their level of security awareness on a scale of 1-7 (from very poor to exceptional). Users were asked to assess themselves at the start of the survey with the same scale, with this idea here to determine if the user still gives the same score having completed the survey in full.

Respondents were then asked to give their opinion on how often security awareness training should be provided by TU Dublin. This was asked to get an overall censes if students thought this should be provided or not.

Finally, the last question asked the respondent if they had any comment to make regarding the survey they have just completed. This allowed the respondent to submit an open-ended response to highlight if any aspects of the survey were incorrect, or if the multiple-choice answers presented to them restricted their answers in a certain way.

3.6 Piloting of the survey

Moser & Kalton (2017) refer to the piloting of a survey as the "dress rehearsal". It is generally done on a small sample of the target population to determine if the questions being asked are phrased correctly and that each question can be understood. Carrying out a survey pilot is crucial in order to achieve research goals and ensure that participants complete the survey (Andrews et al., 2007). It is also helpful in identifying that sufficient responses are available to the participants for each particular question.

Bowden, Fox-Rushby, Nyandieka, & Wanjau (2002) identified that the questions should be placed together as it is expected they will appear in the final survey. Respondents should be given the opportunity to ask for clarification on each question. Bowden et al., (2002) also identified the following questions that should be included in the pilot for this survey. These included the following:

- What they thought about the questions in general
- What they thought about the length of the survey
- If there was any terminology in the questions that they did not understand.

- Whether any questions should not be asked in the survey
- Whether any questions seemed to be strange or unusual

As part of this pilot survey for this research, a total of 21 students were surveyed. This included representation from each of the four colleges within TU Dublin. In addition to students, 11 staff members within TU Dublin also took part in the pilot survey. This ranged from a number of faculty staff (Academic and non-academic staff) as well as members of the IT department. This allowed for expert and non-expert users to assess the questions and allowed for feedback.

A number of these pilot surveys were completed on mobile devices to ensure that the questions were readable, and the use of a smaller screen did not affect the layout of the questions. The initial results of the pilot study identified that a number of the behavioural analysis questions were not phrased in a way that was understandable by a non-tech savvy user.

3.7 Sample Size required

According to Kelley et al. (2003), there is no definitive answer as to what sample size is required for a survey, although larger samples give a better estimate of the population. It is quite rare that everyone asked to participate in a survey will reply (Kelley et al., 2003).

In order to achieve a high number of responses, a link to the survey was e-mailed to all students within TU Dublin. Due to the fact that specific students were not targeted with this email, convenience sampling was used. The relative costs and time used to carry out a convenience sample are small in comparison to probability sampling techniques.

Figure 3-3 below shows Cochran's sample size formula which will be used in this study to calculate the sample size required.

$$n_0 = \frac{(t)^2 * (p)(q)}{(d)^2}$$

Figure 3.3: Cochran's sample size formula

In the above formula, the **t**-value relates to the confidence level. To obtain a confidence level of 99%, the t-value would be set at 2.58. The **p** represents the population split, which is set at 50% (0.5) and **d** is the acceptable margin of error for the proportion being estimated, which in this case is 5% (0.05). Using the above formula, we can estimate that the minimum sample size should be 663. The number of responses obtained in the survey was 752, which exceeded this required figure.

As outlined earlier in this chapter, the student population within the TU Dublin city centre campus is 19,528. Bartlett & Ik (2001) explain that if the sample size calculated using Cochran's (1977) formula exceeds 5% of the population, Cochran's (1977) correction formula should then be used to calculate the final sample size. In this case, the sample size is less than 5% of the population.

3.8 Analysis of Survey platforms

A number of online survey platforms were tested and evaluated for the purpose of running this survey. There was a requirement for the data to be easily exportable to SPSS to allow for the data to be analysed without the need for the data to be converted from a different format.

	Qualtrics	**SurveyMonkey**	**SurveyPlanet**	**Zoho**	**Google Forms**
Price range	High	Low	High	Medium	Free
Data exportable to SPSS	Yes	Yes	No	No	Yes
Supports Question Blocks	Yes	Yes	Yes	Yes	Yes
Limit on Respondents	Unlimited	100 with free version	Unlimited (paid)	150 with free version	Unlimited

Table 3-1: Comparison of online survey platforms

Having assessed the various platforms available, the main criteria that was regarded as essential was the option of separating the questions into different blocks and to allow an unlimited number of responses. All platforms allowed for an unlimited amount of responses, with Google forms being the only platform that provided this at no extra cost. In addition, students within TU Dublin are all provisioned with a G-Suite account and would be familiar with the layout and feel of this online platform. Due to these reasons, Google forms was selected as the platform to host the online survey.

3.9 Statistical tools & methods used

3.9.1 Two-sample t-test

A two-sample t-test is a statistical method that is used to compare if two population means are equal or if there is a significant difference between the two (Snedecor & Cochran, 1989). The data may either be paired or unpaired. For unpaired samples, the sample sizes for the two samples may or may not be equal.

This method was used to determine if there was a significant difference between the mean scores of certain demographic groups within the survey for variables that contained two categorical values. In this case, it was used to compare the quiz score means of the gender of each respondent; male and female as well as the student status of each respondent; full-time or part time student. Figure 3.4 below the formula used to determine the t value in an independent t-test when equal variances are not assumed.

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

Figure 3.4: Formula to determine t-value in an independent t-test

3.9.2 One-factor ANOVA

A one-factor analysis of variance (ANOVA) is used to determine if there are any statistically significant differences between the means of three or more independent

groups (Snedecor & Cochran, 1989). When carrying out this type of method, it is not possible to determine which specific groups are significantly different, only that at least two groups were significantly different.

This method was used to determine if there was a significant difference between the mean scores of demographic groups that contained three or more categories. In this research, the dependent variables were:

1. The mean score of the quiz results.
2. The mean result obtained by each respondent, calculated when their security habits were weighted and scored,

The independent variables related to certain demographic groups which contained three of more categories. The independent variables assessed using this method were as follows:

1. Age range (total of eight different groups)
2. Level of study (total of five different groups)
3. Area of study (total of four different groups)
4. Previous training (total of three different groups)

3.9.3 Chi-square test

A Chi-square test is intended to test how likely it is that an observed distribution is due to chance. It measures how well the observed distribution of data fits with the distribution that is expected, if the variables being analysed are categorical and independent (Maydeu-Olivares & Garcia-Forero, 2010)

A chi-square test was used in this research to establish if the sample population observed in the survey was representative of the actual student population. It was also used to determine if a subset of the sample that stated they did 1) use a device to complete college assignments and 2) opted to declare information relating to their student password habits were representative of the survey sample data.

3.9.4 Statistical tools

A number of statistical tools were assessed as part of this research in order to determine if they could perform the various methods outlined in this section. A license of SPSS is available to use for students within TU Dublin free of charge. SPSS was used to analyse the data obtained in the pilot of the survey, along with some dummy data generated used to fully test the results obtained using both two-sample t-test as well as a one-way ANOVA. SPSS also allows for testing using chi-square. For these reasons, SPSS was selected as the statistical analysis tool to analyse the results of the online survey

3.10 Converting responses to quantitative data

Respondents were asked to provide details relating to their security habits. This included information about their habits relating to their own personal device, awareness of risks with open wireless connections, password habits and data protection. The answers to these questions are all multiple choice. In order to analyse and compare the security habits of respondents within the various demographic groups, a weighting system will be applied to each possible response, with each respondent being assigned a score relating to their security habits.

3.10.1 Device usage habits

Table 3-2 below outlines the questions that will be used for the behavioural analysis, and the corresponding values that will be applied to each response. A total of nine questions are outlined below.

Device Usage		
Q DU1. Is your device password protected?		
	Response	Weighted value
	Yes	1
	No	0
	I'm not sure	0
Q DU2. Do you have Anti-virus / Anti-malware installed on the device?		
	Response	Weighted value
	Yes	1
	No	0
	I'm not sure	0
Q DU3. How often do you install OS updates?		
	Response	Weighted value

As soon as I am prompted	1
My machine is set to automatically update itself	1
I don't install updates	0
Not sure	0
I don’t know what an OS update is	0

Q DU4. How often do you install software updates on your device (e.g. Web Browsers, Office Products)?

Response	Weighted value
As soon as they are released	1
I only update software if it starts causing problems	0
I don't normally update software on my machine	0
I’m not sure	0

Q DU5. Do you have a firewall enabled on the device?

Response	Weighted value
Yes	1
No	0
Not sure	0

Q DU6. Do you backup data on your device?

Response	Weighted value
Yes	1
No	0

Q DU7-1. Do you allow other users to use your device?
&
Q DU7-2. Do you allow your web browser (such as Google Chrome) to store your passwords?

Response Q DU 8-1	Response Q DU 8-2	Weighted value
No	/	1
Yes/Maybe	No	0
Yes/Maybe	Yes	-1

Q DU8. Is the account you primarily use on your device an admin user?

Response	Weighted value
Yes	0
No	1
Not sure	0

Q DU9 Are you aware that using appropriate tools, a hacker could intercept your wireless traffic if you are using an open/insecure network?

Response	Weighted value
Yes	1
No	0
I don’t care	0

Table 3-2: Device usage responses converted to numerical values

In relation to encryption, some Windows operating systems have a built-in encryption tool called BitLocker. Due to the limitations with certain versions of Microsoft Windows, BitLocker is not included with all versions. For example, only the Enterprise and Ultimate versions of Windows 7 and Windows Vista include Bitlocker (Casey,

Fellows, Geiger, & Stellatos, 2011), whereas all Apple laptop devices, since Mac OS 10.3, include FileVault, which is Apple's built in full disk encryption solution. For this reason, the question relating to whether or not the respondent's device was encrypted was not included in this scoring.

3.10.2Password habits

Table 3-3 below outlines the questions from the survey that will be used to assess the password hygiene of each respondent along with the corresponding weightings that will be applied to each response.

Password hygiene

Q PW1. How often do you change the password on your student account?

Response	Weighted value
Never	0
Once	1
Regularly	2
Not sure	0

Q PW2. Thinking of the password you use for your student account, how long is this password?

Response	Weighted value
8 characters	0
9-11 characters	1
Longer than 12	2
Rather not say	N/A

Q PW3. Have you ever used the same password on multiple sites?

Response	Weighted value
Yes	0
No	1
Rather not say	N/A

Q PW4. In relation to your online accounts (social media, email etc.), do you use a 3rd party password manager to store your passwords?

Response	Weighted value
Yes	1
No	0
I'm not sure what a password manager is	0

Q PW5. Do you know what Two-Factor Authentication is (also known as Multi Factor Authentication) and have you implemented this on any of your online accounts where it is offered?

Response	Weighted value
Yes, and I have implemented it on all or some of my online accounts	1
Yes, but I have not implemented it	0
No, I don't know what it is	0

Table 3-3: Password Hygiene responses converted to numerical values

Due to the recommendations outlined by Grassi et al. (2017), students that opted to use a longer password were awarded a higher score. Students that regularly changed their password were also awarded a higher score than those who only changed it once or never changed it.

3.11 Summary

In this chapter, the various demographic groups were described in detail, with the reasons behind selecting each categorical data to represent this data explained. Reasoning of why a survey was used were discussed, along with details of the pilot study, which was representative from students within the four colleges, along with academic staff and IT to get feedback from all user types.

Using Cochran's sample size formula, it was determined that the minimum number of respondents required to give a 99% confidence level was 663. Various statistical tools were assessed, with the one chosen to analyse the data being SPSS. Finally, the security behaviour scoring of respondents was outlined in two table; one for device security, the other for their password habits.

4 RESULTS & OBSERVATIONS

4.1 Introduction

The objective of using a survey was to evaluate a sample of the student population to determine their level of information security awareness. In this chapter, the respondents quiz scores will be assessed, and a number of methods will be used to determine if there is a significant difference between the various demographic groups when the mean scores obtained in the quiz are evaluated.

Respondents' behaviours will also be assessed using the weighting and scoring outlines in the design and methodology section of this document. Once these have been calculated for each respondent, a similar exercise will be carried out to determine if there is a significant difference between these mean scores when the demographic groups are compared. Finally, the mean score of both the quiz and the respondents behaviours will be compared for respondents who have participated in training or not, to determine if there is significant difference between these groups and to determine if ISA training has any impact on a user's awareness of security best practices and their own security habits.

4.2 Survey Responses

An email inviting all students to participate in the survey was sent out on Monday 1st of April. The survey was left open for a total of ten days. A total of 752 surveys were fully completed. Each question within the survey was marked as mandatory, excluding the comment field at the end of the survey. This ensured that all questions asked were answered by each respondent. Any surveys which were not fully completed were not recorded within Google Forms. Figure 4-1 below gives an outline of the survey response rates over the course of the ten days. The majority of the surveys were completed on the first day, with a significant drop off after the third day.

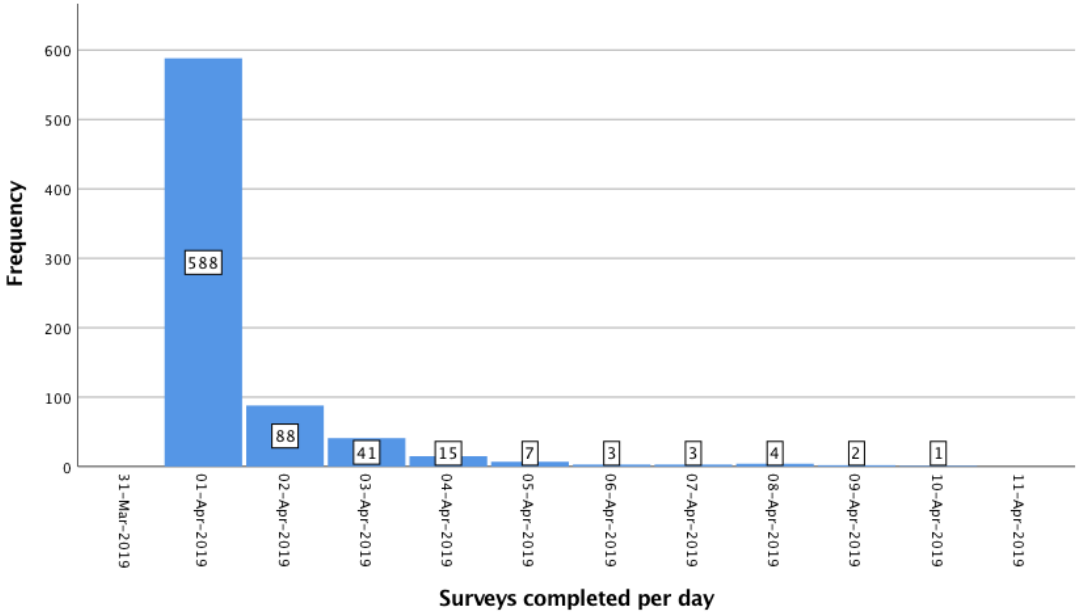


Figure 4-1: Breakdown of number of surveys completed

4.3 Demographic Survey

As described in the Design and Methodology section of this document, the first part of the survey collected demographic information from each respondent. The information gathered in this part of the survey was deemed relevant based on previous surveys carried out in this area of research

As discussed in the design and methodology section of this document, the sampling method used in this survey was convenience sampling, as there was not enough time to use probability sampling methods, such as simple random sampling or stratified random sampling. A chi-square test was carried out to determine how significantly different the sample obtained varied from the actual population. These figures can be found in Appendix B. Due to the fact that probability sampling was not used, it was not expected that a chi square "goodness of fit" test would determine if the respondents who completed the survey were a good representation of the population.

4.3.1 Gender

The bar chart below related to the breakdown of male and female respondents. Although a total of 752 respondents completed the survey, a small number (8) selected the option

of not specifying their gender or selected other for their gender type. A breakdown of these figures can be seen below in Figure 4-2Figure 4-2

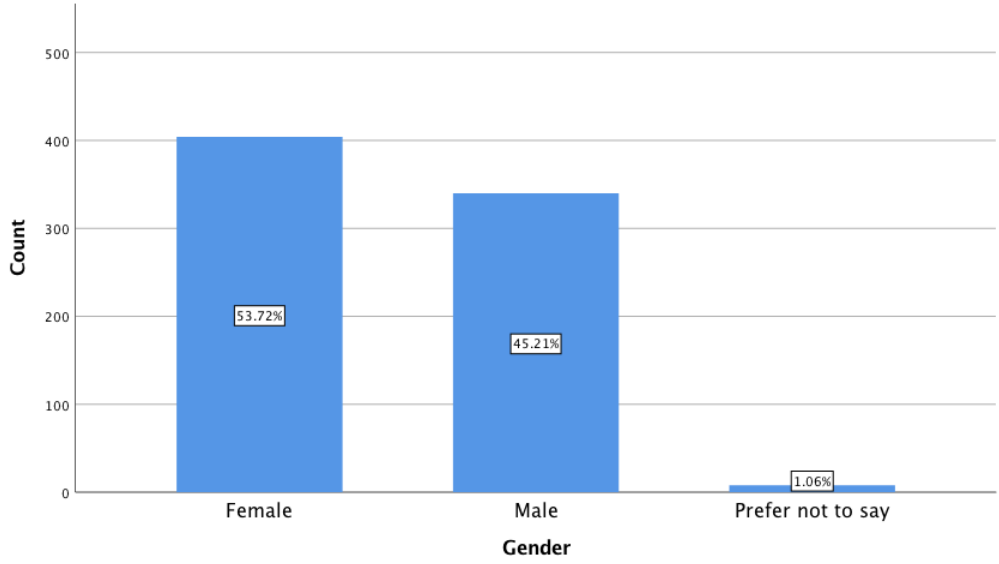


Figure 4-2: Breakdown of Gender

4.3.2 Age Distribution

The age distribution of respondents is outlined in Figure 4-3 below. Similar to gender, age is another factor that was used to compare demographics in previous studies. The chart below shows a reduction in the number of responses as the age group increases. It was expected that based on the statistical information available from the Higher Education Authority of Ireland, the age range of 20-21 is the most represented group between both full-time[2] and part-time[3] students.

[2] http://hea.ie/assets/uploads/2018/09/Full-Time-Enrolments-by-Gender-and-Age-2017-18.xlsx

[3] http://hea.ie/assets/uploads/2018/09/Part-time-Enrollments-by-Gender-and-Age-2017-18.xlsx

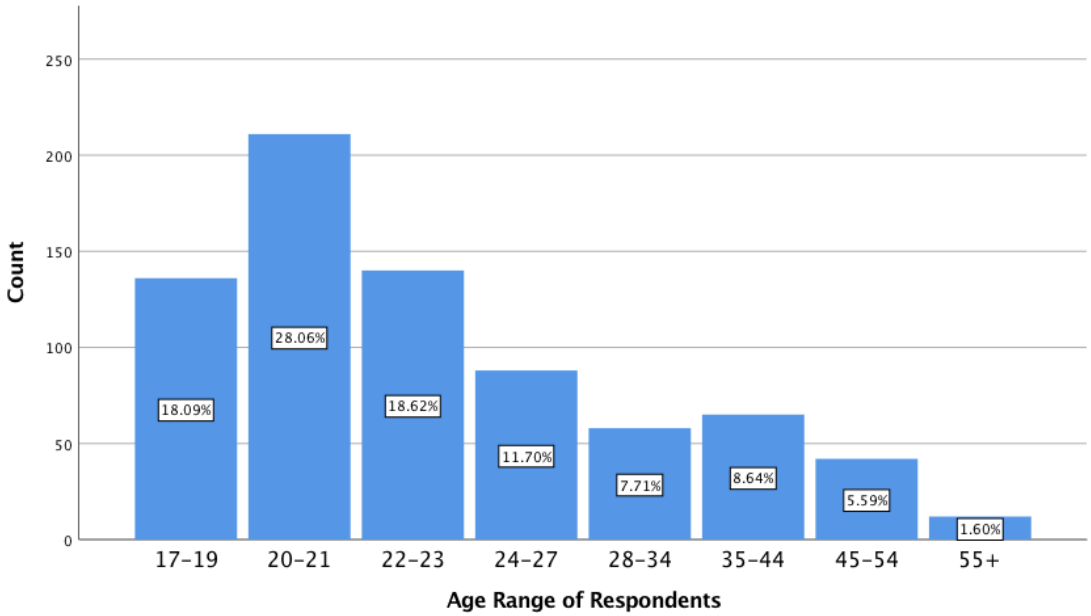


Figure 4-3: Age breakdown of survey respondents

4.3.3 Education

The survey contained three questions in relation to the level of study of the respondent and an additional question asking if they had every participated in Information Security awareness training in the past.

In relation to the level of education, this information was categorized into the level of study the student was currently at, the area of study, which was based on the college the student was currently enrolled in and the status of the study; whether they were a full-time or part-time student.

Figure 4-4 below identifies the area of study each respondent is based in based on their gender. The highest number of surveys were completed by students based in the College of Science and Health (31.51%), with the lowest response rate coming from the College of Arts and Tourism.

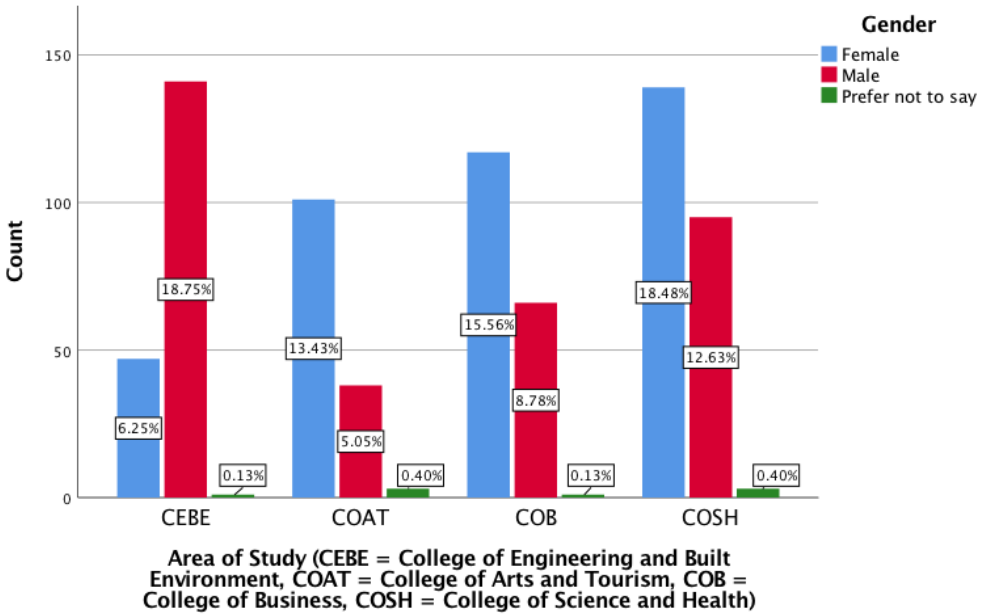


Figure 4-4: Area of study breakdown by gender

Using a clustered bar chart, it is possible to give a breakdown of full-time and part-time students within each of the four colleges. This can be seen in Figure 4-5

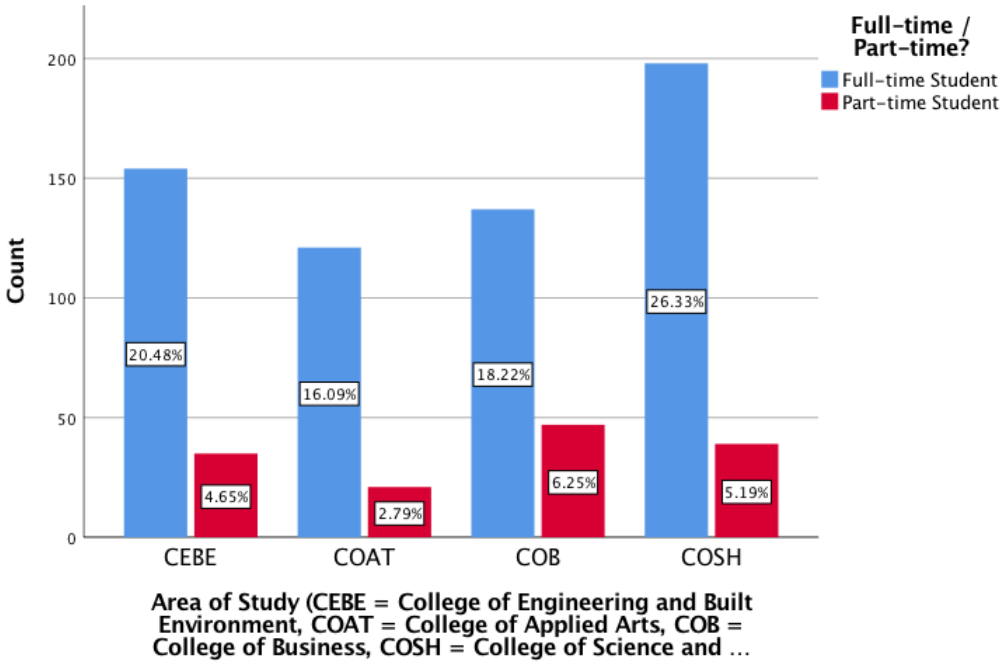


Figure 4-5: Breakdown of full-time and part-time students per college

In relation to the current level of study that the student is currently at, the clustered bar chart below gives a breakdown of the level of study of each respondent per college.

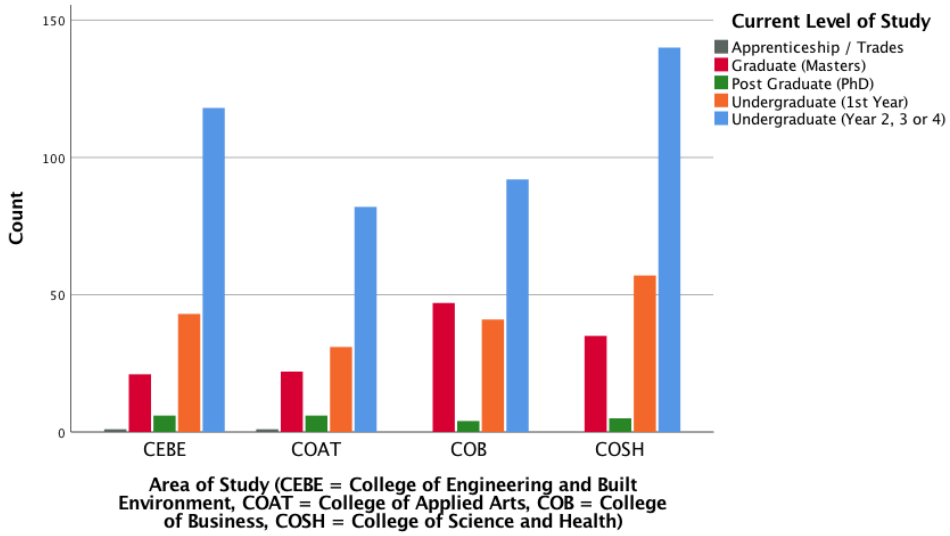


Figure 4-6: Breakdown of respondents' level of study per college

Respondents were also asked if they had participated in information security awareness or cyber security training in the past. As can be seen from Figure 4-7 below, more than 84% of respondents answered "No" or that they were "Not sure". Just under 12% of respondents had participated in this type of training within the past 2 years.

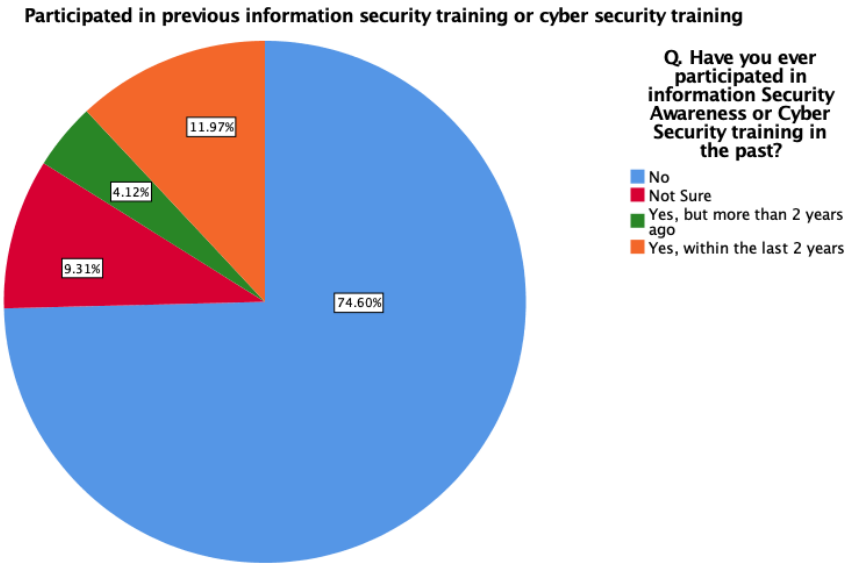


Figure 4-7: Breakdown of respondents that had participated in information security training in the past

4.4 IT Competency & Security awareness

Each respondent was asked at the start of the survey to rate their level of IT competency. This was set on a Likert scale that ranged from 1-7, with 1 = very poor, 2= poor, 3= fair, 4=good, 5=very good, 6=excellent and 7 being exceptional.

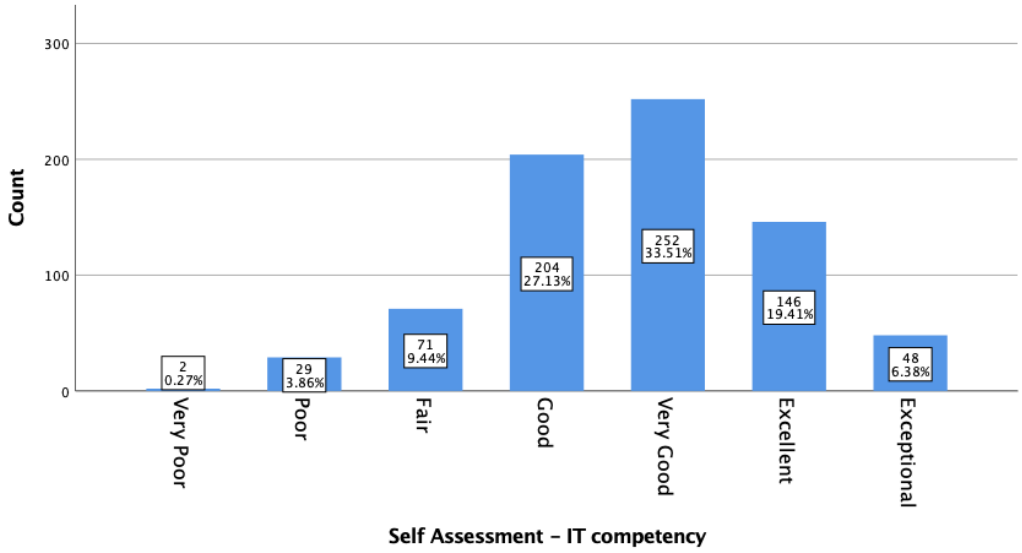


Figure 4-8: Self-assessment of IT competency

Respondents were then asked to do a self-assessment on their IT security awareness using a similar scale. A high number of users claimed to have a higher IT competency level, with more than 59.3% stating that their IT competency level was regarded as "very good" or better. In comparison, only 34.9% of respondents assessed that their IT security awareness was at the level of very good or higher. More worryingly, 37.6% of respondents claimed that their IT Security awareness was regarded as either "very poor", "poor" or "fair".

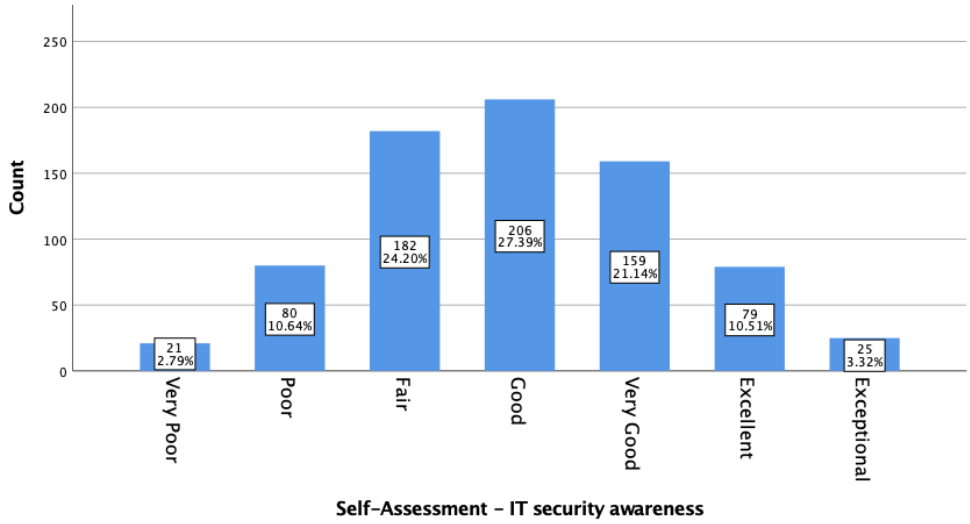


Figure 4-9: Self-assessment of IT security awareness

A clustered bar chart was used to show the breakdown of the self-assessment in IT security awareness based on the area of study the respondent stated that they were enrolled with. As can be seen in Figure 4-10 below, a large number of students based in the college of science and health stated that their level of IT security awareness was regarded as "Good" or higher.

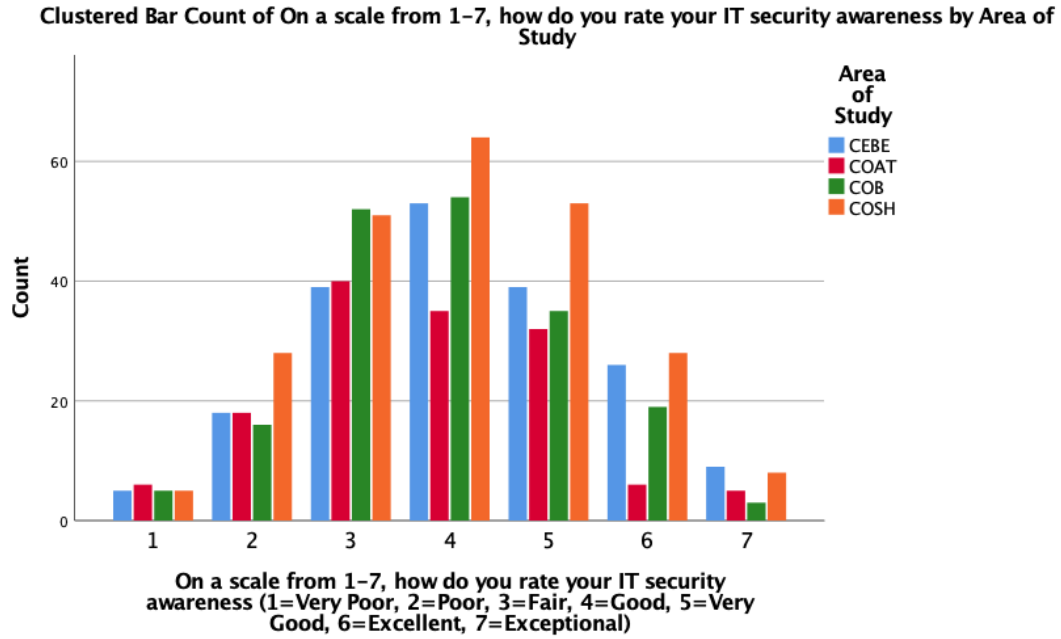


Figure 4-10: Self-assessment of IT security awareness by area of study

4.5 Prior security breaches

After respondents were asked to assess their level of security awareness and IT competency, they were then asked if they had ever been involved in a security breach, particularly in relation to having their email account, online shopping account, online banking account or any of their social media accounts compromised. Over 41% of respondents had claimed that one of their account had in fact being breached.

4.6 Personal Device Usage

As outlined in the design and methodology section of this document, the student helpdesk was asked to take note of the type of device that was used by each student that had called into the helpdesk looking for assistance. Over a two-week period, the student helpdesk noted that the majority of devices (76%) used by students was in fact a PC laptop running Microsoft Windows, with 23% of device being Apple Mac Book device. Figure 4-11 below gives an overview of the device break down for each respondent.

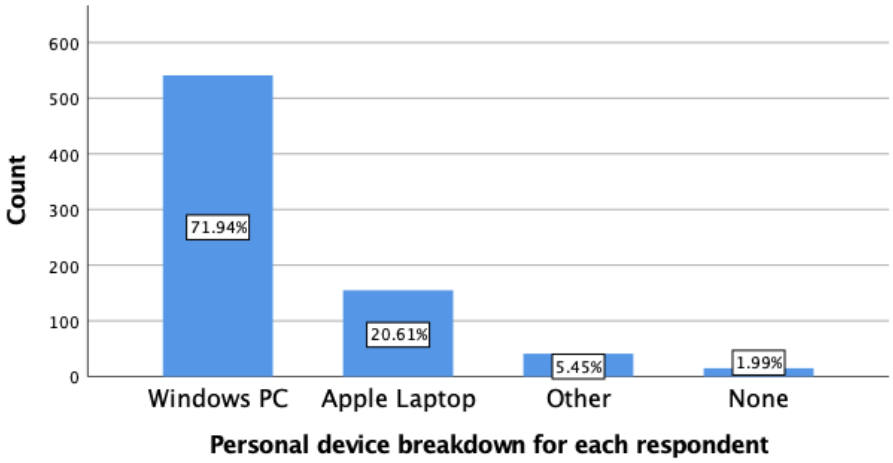


Figure 4-11: Breakdown of devices used by each respondent

The number of PC and Mac devices recorded in the survey are representative of the figures observed by the helpdesk over the two-week period.

4.6.1 Windows PC Laptop Users

A total of 541 respondents stated that they used a Windows PC laptop to complete university assignments. Figure 4-12 shows the percentage breakdown of the various Microsoft Windows operating systems that are used by each respondent.

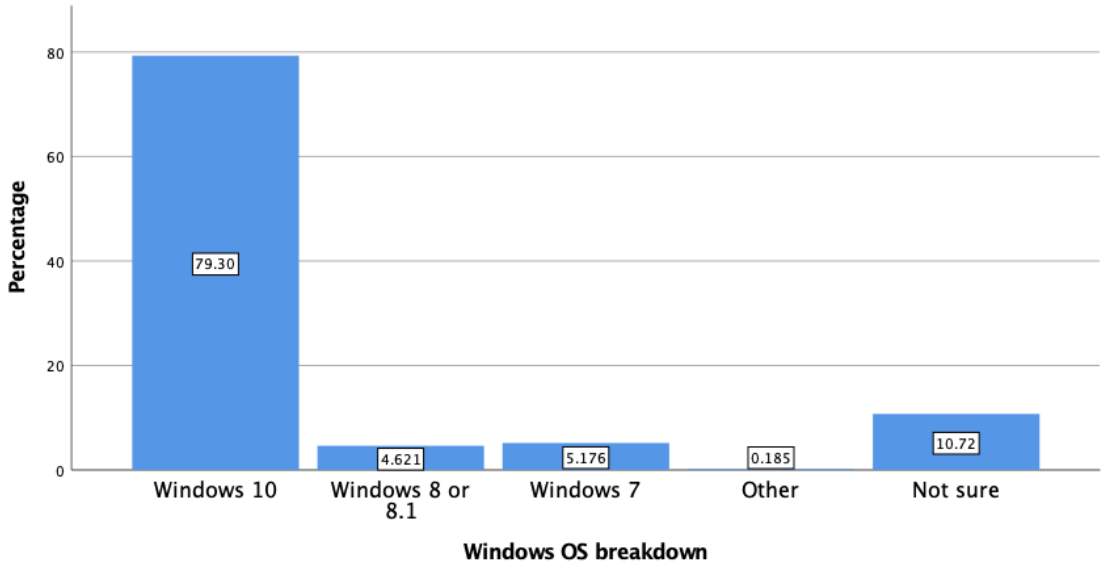


Figure 4-12: Percentage breakdown of Windows operating systems used by respondents

Only one respondent indicated that they were using a windows OS that was not listed on the survey, with over 10% not sure as to which Windows operating system they were using.

4.6.2 Apple Laptop Users

A total of 155 respondents stated that they used an Apple Mac Laptop as their primary device for completing college assignments. Figure 4-13 below gives the percentage breakdown of the different Mac OS versions running on each device. Surprisingly, over 32% of users that state they use an Apple Mac Laptop were unsure of the version of operating system on their device. 8.39% of users were using an unsupported version of Mac OS, meaning that security updates are no longer available for these versions. Nearly 60% of users were using a version that was still supported by Apple.

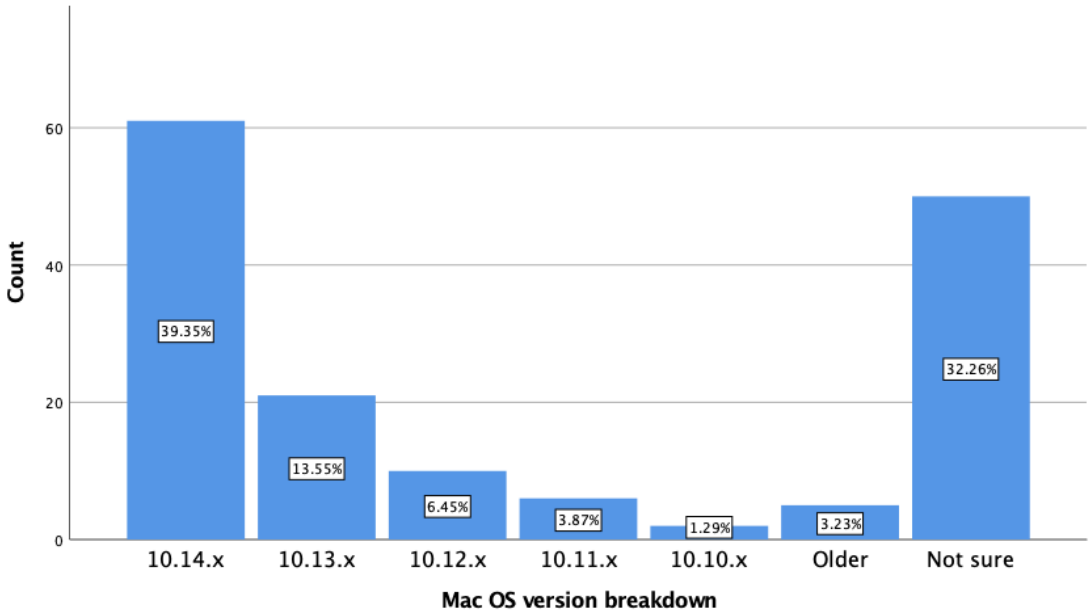


Figure 4-13: Percentage breakdown of Mac OS versions used by each respondent

4.7 Device Encryption

Respondents who stated they used a Windows PC laptop or a Mac Laptop as their primary device were asked if their device was encrypted. Windows has a built-in encryption tool known as BitLocker, but not all versions of Windows are bundled with this (Casey et al., 2011). Due to this, respondents were not questioned specifically if they had Bitlocker enabled, but just if the device had been encrypted, with the multiple-choice options being "Yes", "No" or "I'm not sure".

Mac OS devices have a built-in encryption tool known as FileVault, which is bundled with every version of Mac OS since version 10.3, which was released in 2003 (Joyce, Powers, & Adelstein, 2008) . Due to this, respondents were asked if FileVault was enabled on their device, with the multiple-choice options being "Yes", "No", "I'm not sure" or "I've never heard of FileVault". Due to the differences in the possible answers, the results are presented in two separate graphs below.

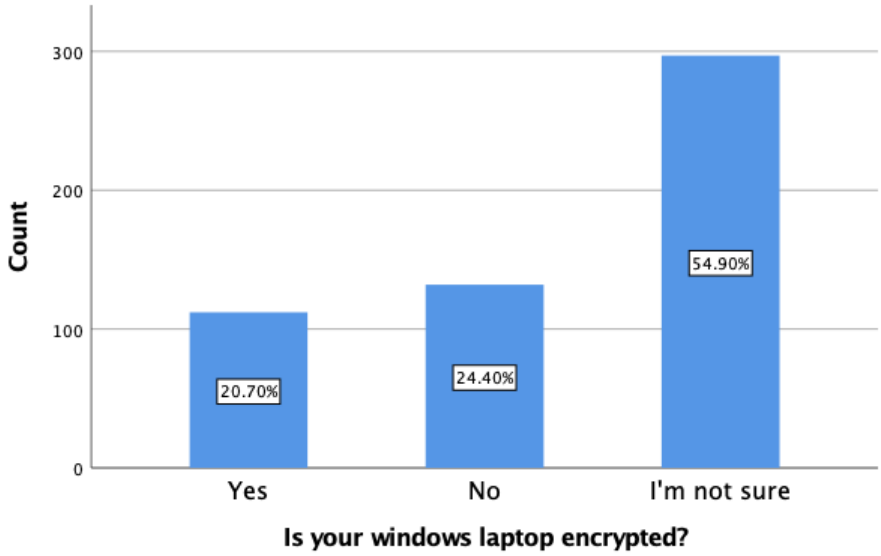


Figure 4-14: Percentage of Windows laptops encrypted

As can be seen in Figure 4-14, only 20.7% of users of Windows laptop devices have encryption enabled on the device. This may be to do with BitLocker not being bundled with every version of Windows.

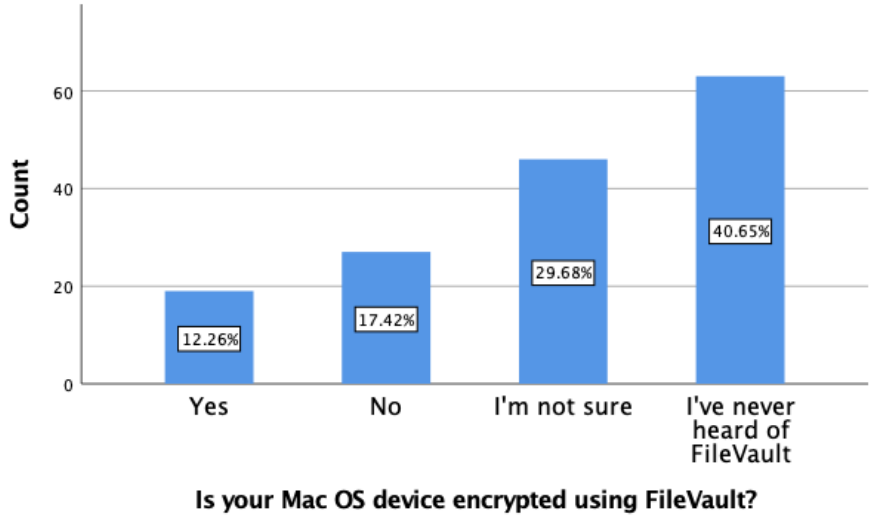


Figure 4-15: Percentage of Mac OS devices encrypted

Similiary, only 12.26% of Mac OS users have encryption enabled on the device, with nearly 30% of these users not sure if it was enabled or not. Surprisingly, over 40% of users have never heard of FileVault.

4.8 Password hygiene - student account

Respondents were asked a series of questions in relation to the password on their student account. When an account is created for a student in TU Dublin, the account is created with a random password, which students are advised to change as soon as they receive it. Due to limitations within TU Dublin on how this password is distributed to students, it is not a mandatory requirement for each student to change their password when the account is created. As well as this, passwords do not expire, meaning it is not a requirement for students to change their password at regular intervals. Due to this password policy, students could use the same password for the duration of their course, which may be at least four years in length.

The first question in relation to password hygiene asked each respondent how often they had changed their student account password, with the option being Never, once, regularly or not sure. As can be seen in Figure 4-16, 49% of respondents stated that they had never changed their password since they had received their credentials, with over 41% stating that they had only changed the password the once.

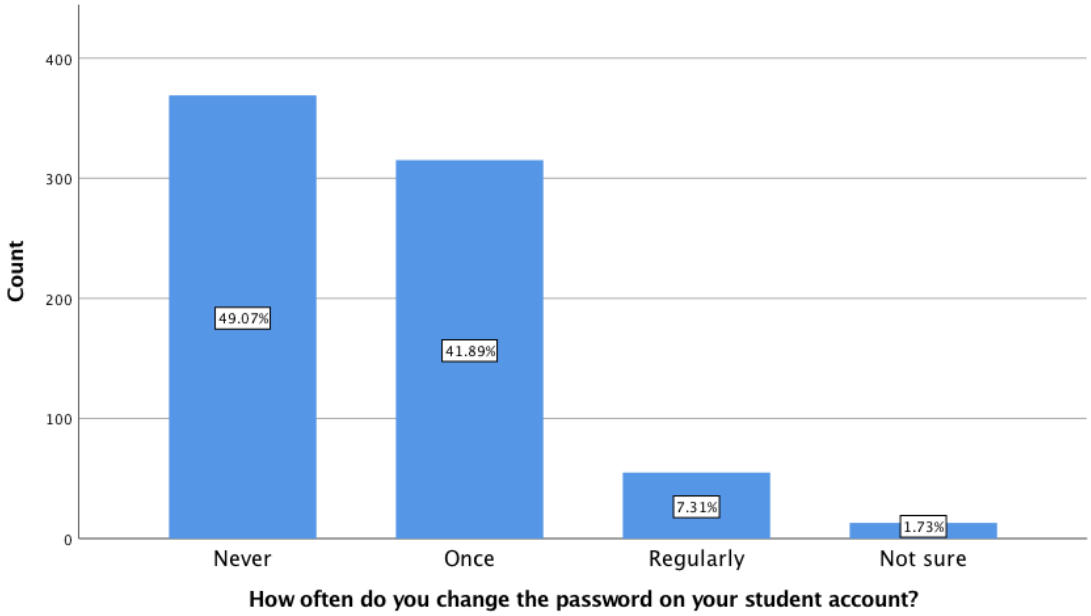


Figure 4-16: Breakdown of how often respondents' change their student account password

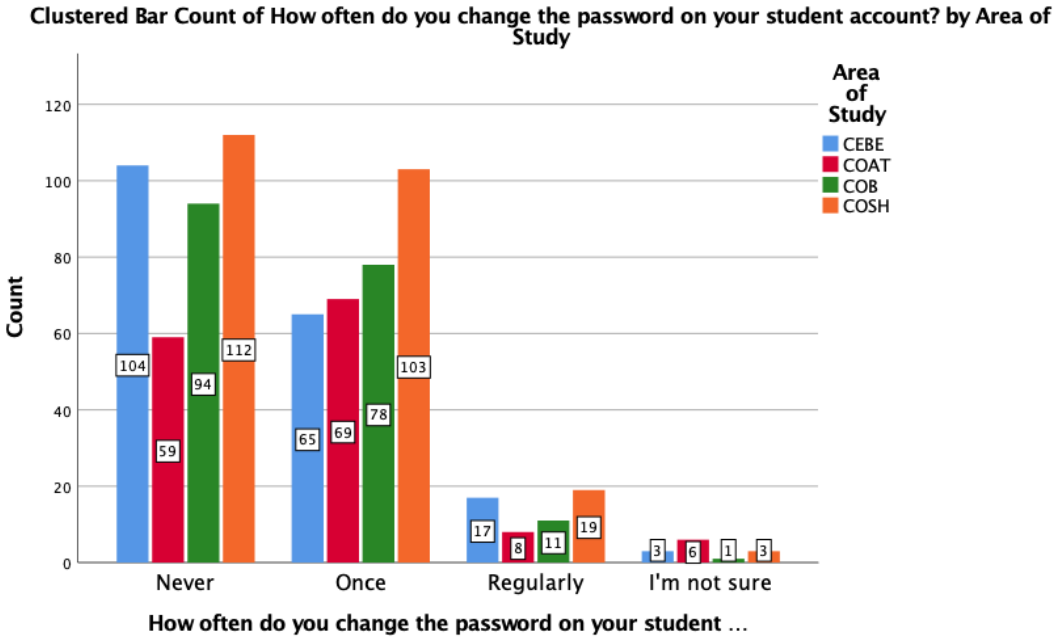


Figure 4-17: Breakdown of how often respondents' change their student account password by area of study

4.8.1 Password Length

In relation to password length, respondents were given the option of stating how long the password was. The 49% of respondents stating that they had never changed their passwords were removed from this analysis, as if they have never changed their password, the password would be the same as it was set by the University, meaning that the student did not create the password. A total of 383 respondents had stated they had changed their password at least once, regularly or that they were not sure. Figure 4-18 below gives the breakdown of password length for each student account where the password has been changed at least once.

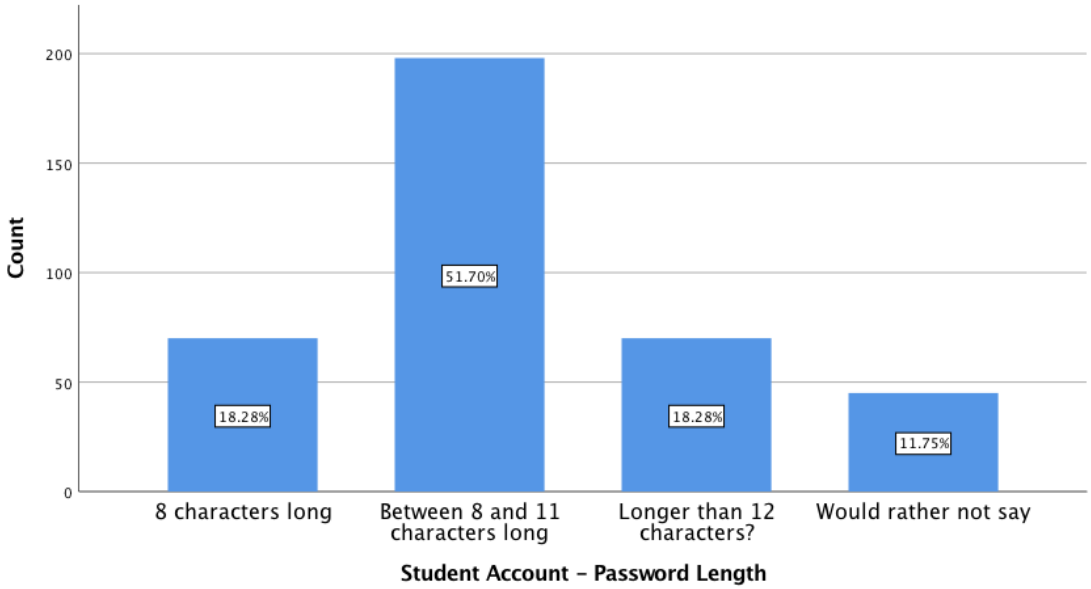


Figure 4-18: Breakdown of respondents' password length for their student account

As can be seen, only 18.28% of respondents that gave information about this stated that their password was 12 characters in length or more. As discussed in the literature, passwords that are too short can yield to brute force attacks and dictionary attacks by using words and commonly chosen passwords (Grassi et al., 2017)

By using a clustered bar chart, we can show the breakdown of this data by male and female respondents. Out of the 383 respondents that stated they had changed their student account password, 208 of these were female, 169 were male and 6 did not state their gender. A large proportion of female respondents stated that their password was exactly 8 characters in length, with a higher number of male respondents stating that their password was longer than 12 characters. Further charts relating to education can be found in Appendix C

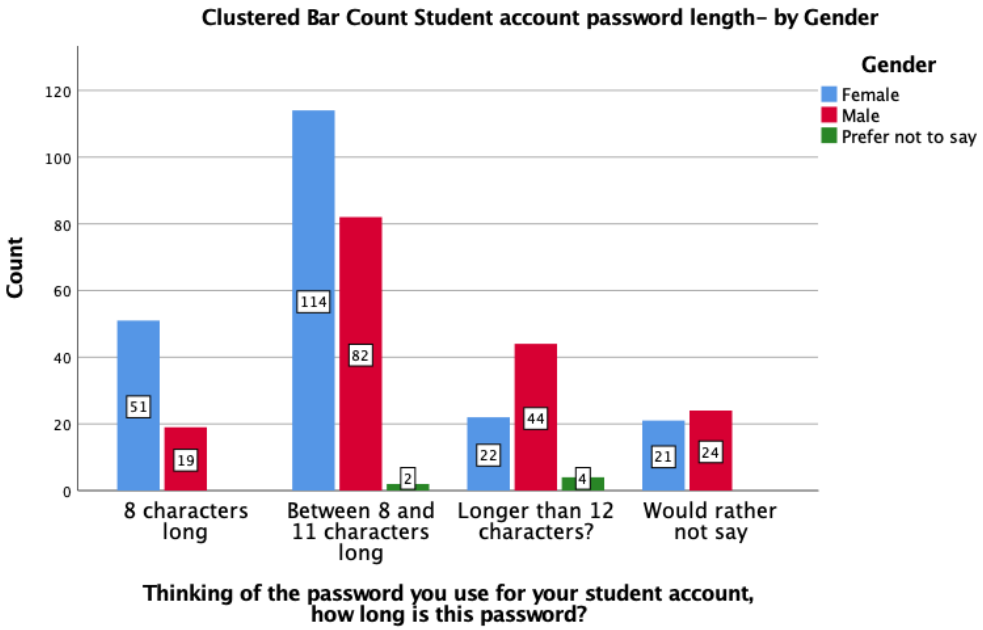


Figure 4-19: Breakdown of student password length by gender

4.8.2 Password Complexity

Respondents were also asked to give details of their password complexity for their student account. As with password length, only students that had stated they had changed their password at least once were included in these figures. It was established that TU Dublin does not implement a password complexity policy, meaning passwords can contain any type of character, and do not need a combination of a certain type of character for the for the password to be regarded as a valid password.

For this question, respondents were advised that complexity was defined as how many of the following types of characters the password contained from the following sections; (1) Lowercase letters, (2) Uppercase letters (3) Numbers, (4) Special Characters.

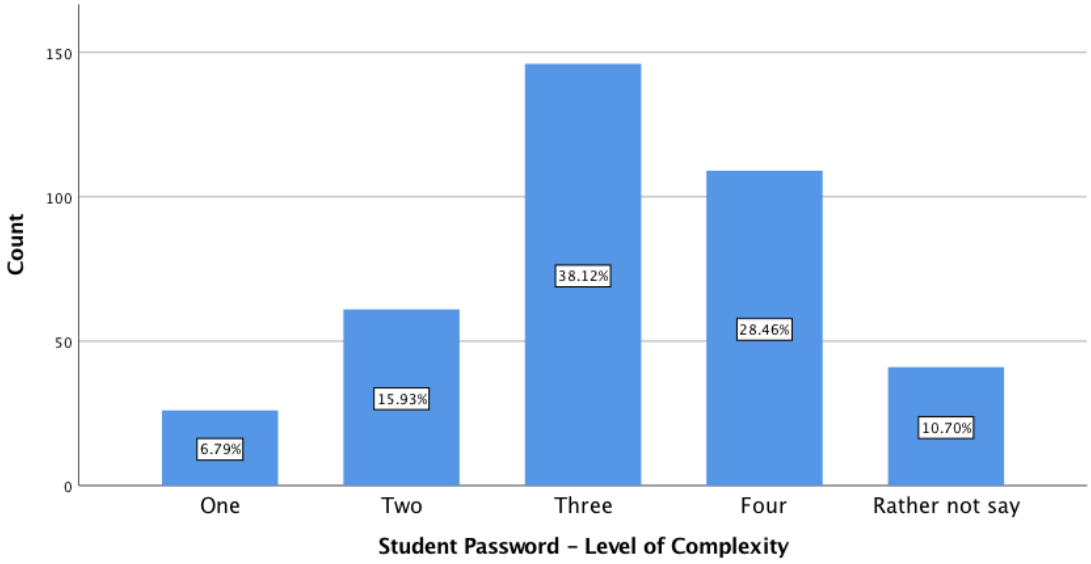


Figure 4-20: Breakdown of respondents' password complexity on their students account

As can be seen in Figure 4-20 above, 6.79% of respondents that gave an answer to this questions stated that they used only one type of these character types in their passwords, with over 66% using at least three types of these character types in their password.

When a comparison was made between male and female respondents in terms of complexity, there was no significant difference between the two groups in relation to password complexity. Likewise, there was also no noticeable difference with the use of password complexity when respondents within the four colleges were compared. Please see Appendix C for this breakdown of college, gender and status of student

4.9 Password hygiene - other accounts

Respondents were asked a series of questions relating to their password habits in relation to other accounts they used, such as social media accounts and other email accounts. The first of these questions asked how often they would generally change their password on these types of accounts. Figure 4-21 below gives the breakdown of answered submitted by each respondent.

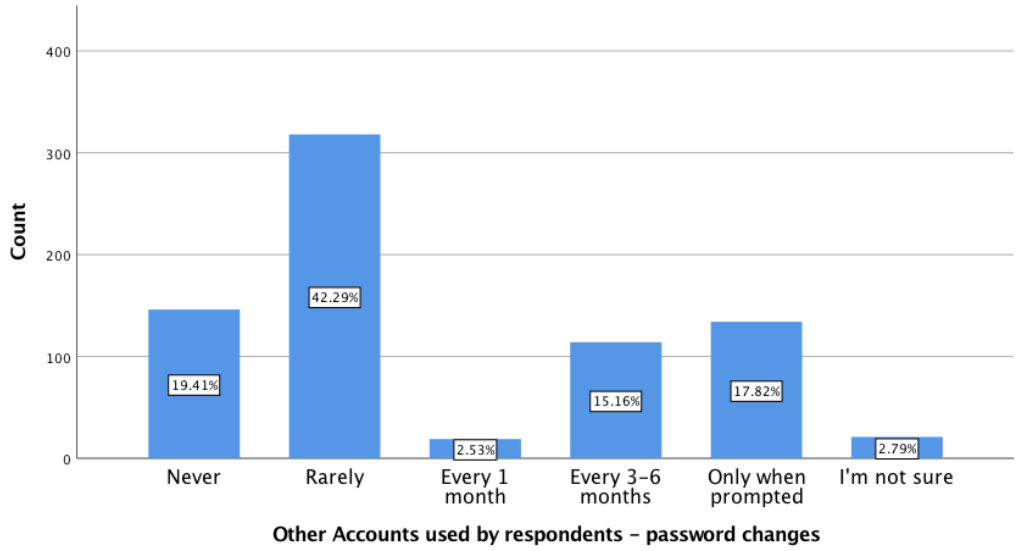


Figure 4-21: Breakdown of how often respondents' changed their own personal account passwords

As we can see, 19.41% of respondents stated that they never change their password, with 42.29 stating that they rarely do. Respondents were also asked if they have ever used the same password on multiple websites. Over 44% of respondents stated that they generally use the same password for all accounts, with 34% stating that they use the same password on some of their accounts.

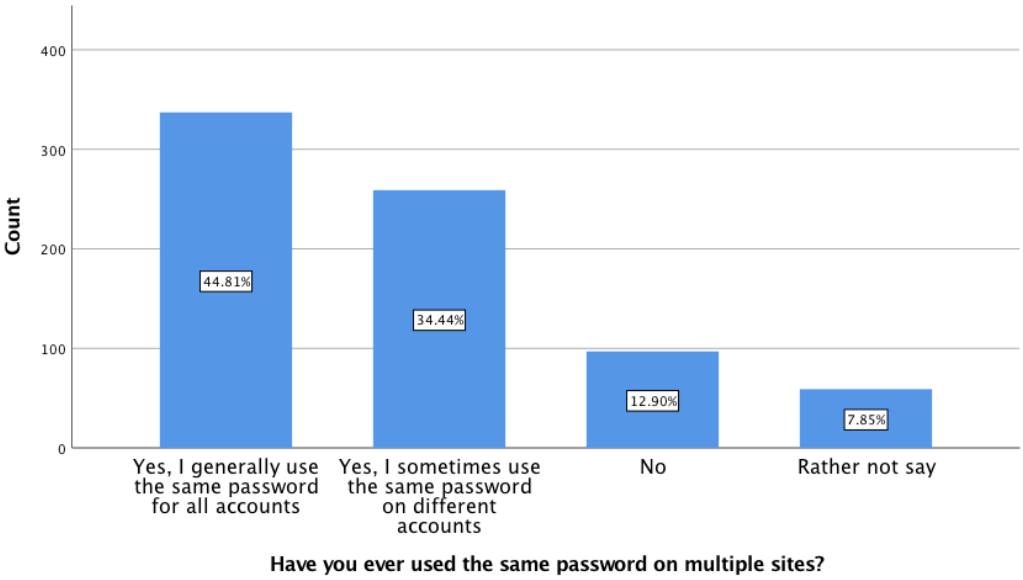


Figure 4-22: Breakdown on respondents' password re-use on personal accounts

A clustered bar chart was used to show the breakdown of both male and female respondents in relation to the use of the same password on multiple sites.

As can be seen in Figure 4-23 below, a higher proportion of female respondents stated that they done this all of the time, with a higher number of male respondents stated that they use a different password for each site.

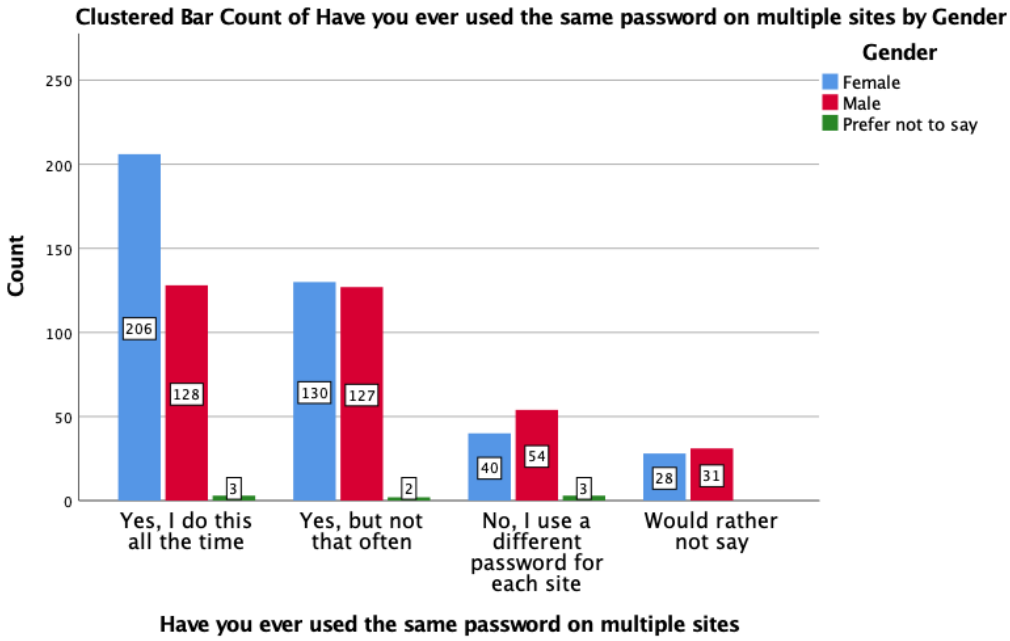


Figure 4-23: Breakdown of respondents' password re-use on personal accounts by gender

A second clustered bar chart is used below to show the breakdown by age in relation to the use of the same password on multiple sites. As can be seen in Figure 4-24 below, students in the age range of 17-21 are more prone to use the same password on multiple sites.

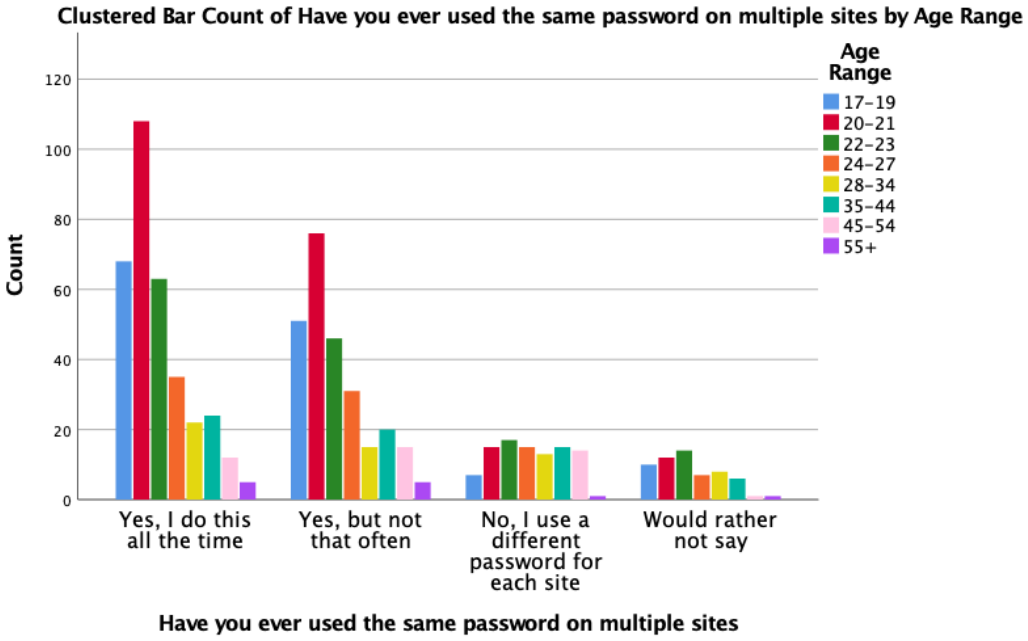


Figure 4-24: Breakdown of respondents' use of password re-use on personal accounts by age

Additional data relating to password habits can be viewed in Appendix C

4.9.1 Password Managers

As outlined in the literature review, password managers were created to relieve password fatigue and facilitate better password quality and a reduction in password re-use across multiple site (McCarney, Barrera, Clark, Chiasson, & van Oorschot, 2012). Respondents were asked if they used a third-party password manager in order to store their passwords for their social media or email accounts. As can be seen in Figure 4-25 below, only 23.94% of respondents claimed to use one of these services.

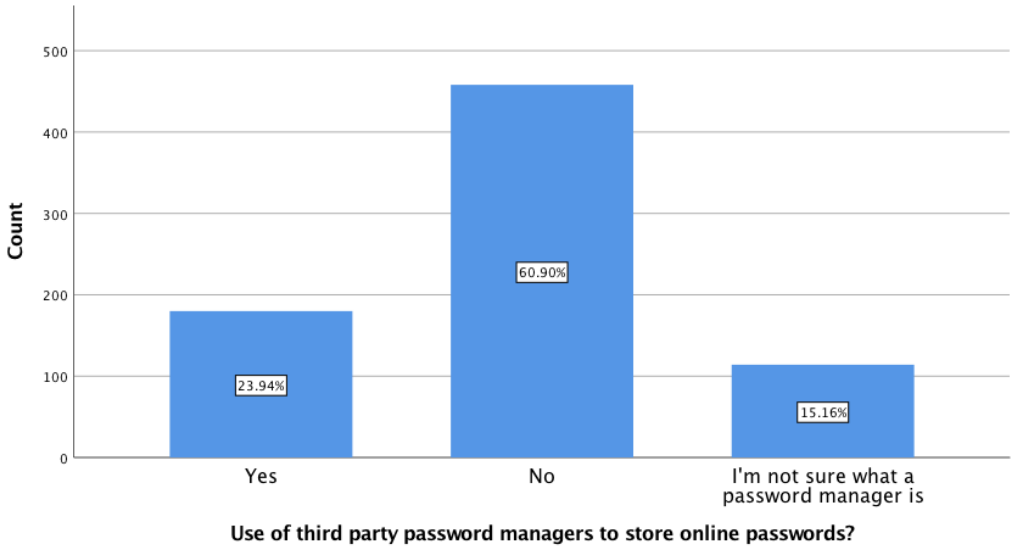


Figure 4-25: Breakdown of respondent's that use a third party password manager

A clustered bar chart was used to show the breakdown of male and female students who used a third-party password manager in order to store their passwords

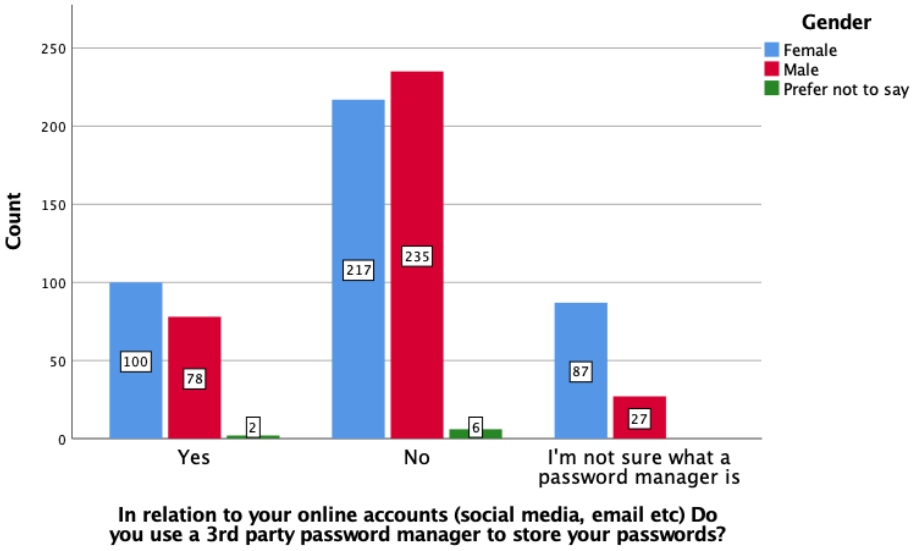


Figure 4-26: Breakdown of respondents' use of password managers by gender

As can be seen in Figure 4-26 above, although the amount of male and female users who did not use a password manager were very similar (217 female respondents compared with 235 male respondents), a higher amount of female respondents stated that they were not sure what a password manager was, with a ratio of just over 3:1.

4.9.2 Two-factor authentication / MFA

Another recent security trend is the implementation of multi-factor authentication, which is available on a wide range of services. Respondents were asked if they knew what MFA was and if so, if they had implemented this on all or some of their online accounts. Figure 4-27 gives the breakdown of these results. Just under half of the students surveyed (48.27%) were aware of MFA and had implemented on some or all of their accounts.

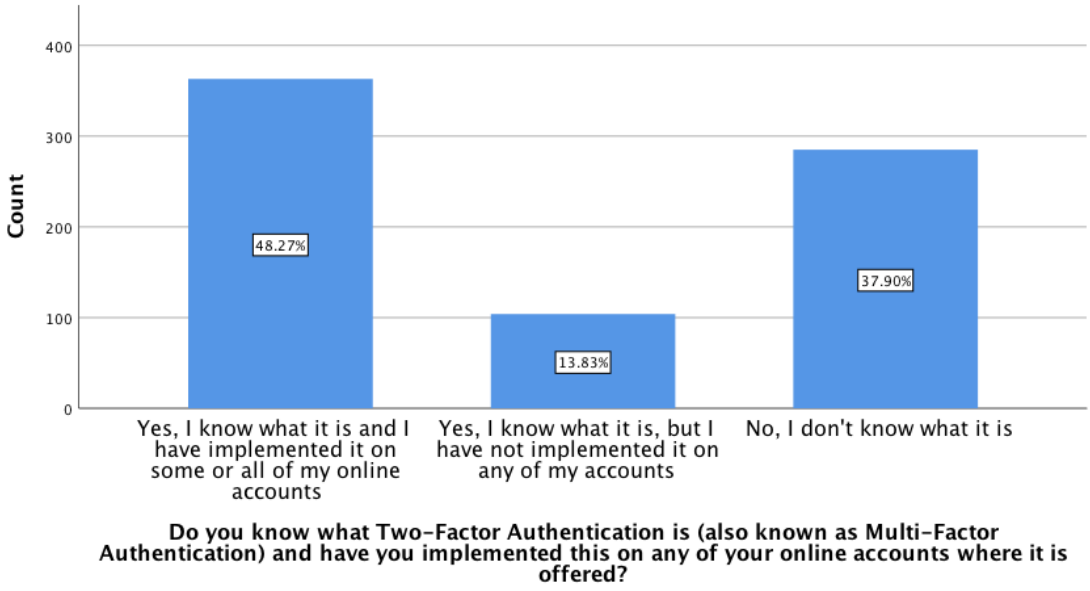


Figure 4-27: Breakdown of respondents' use of MFA

A clustered bar chart was used to show the breakdown of this data in relation to male and female respondents. As can be seen in Figure 4-28 below, a significantly higher amount of the respondents that stated they did not know what two-factor authentication was were female.

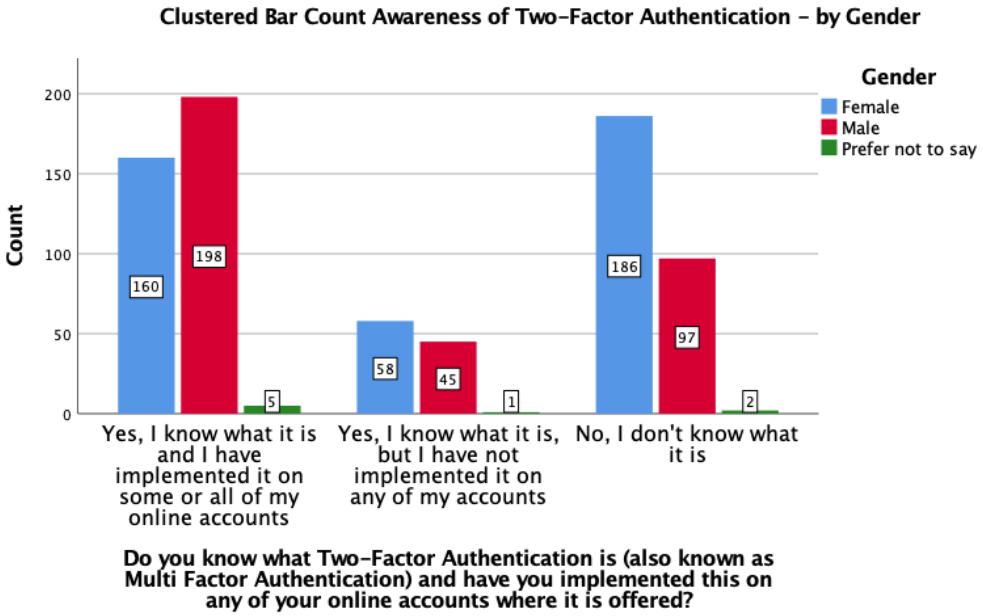


Figure 4-28: Breakdown of respondents' use of MFA by gender

4.10 Insecure wireless connections

Respondents were asked a series of questions relating to wireless technologies, and in particular, their own behaviour when connecting to open/insecure wireless connections. As can be seen in Figure 4-29 below, over 76% of respondents said they had connected to an open / insecure wireless connection from their own laptop or mobile device in the past.

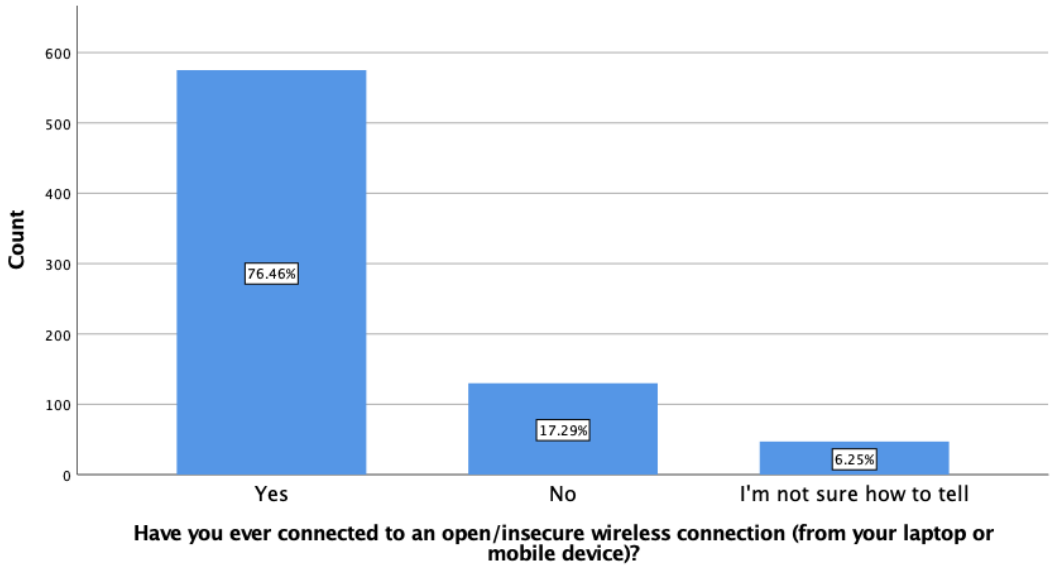


Figure 4-29: Breakdown of respondents' use of insecure network access

Respondents who answered yes to this question were then asked if they had ever logged into their online banking or sent an email over this type of open connection. Out of the 575 respondents (76.46%) that answered yes to the previous question, 37.2% of these stated that they had either accessed their online banking or sent an email over this insecure connection.

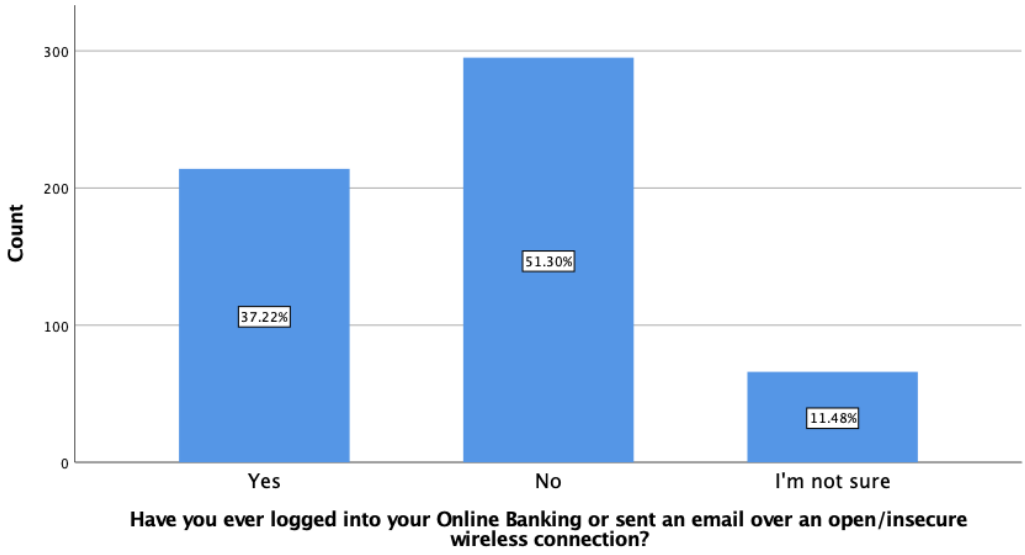


Figure 4-30: Breakdown of respondents' use of accessing online banking or email over an insecure connection

The final question asked in relation to open / insecure wireless connections was if the respondent was aware that by connecting to this type of wireless connection, a hacker could potentially intercept their network traffic. Although a high percentage stated that they were aware of this, just over 22% stated that they were not.

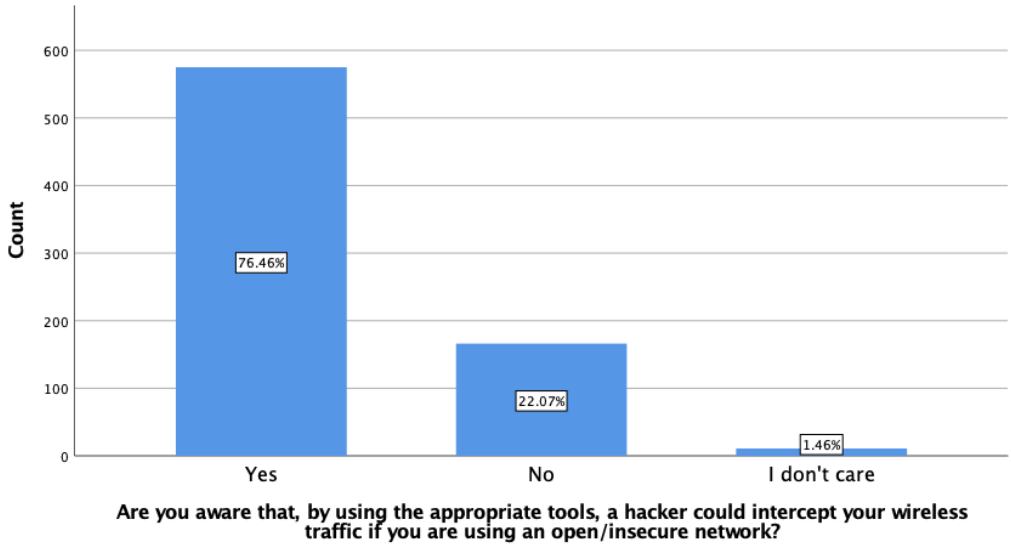


Figure 4-31: Breakdown of respondents' awareness of hacker intercepting traffic over open wireless network

A clustered bar chart was used below in Figure 4-32 to show the breakdown of Male and Female respondents that were aware of the risks of using an insecure wireless connection. As can be seen in the figures below, more than double the number of respondents who answered "No" to this question were Female, in comparison to the respondents that answered "Yes" to this question.

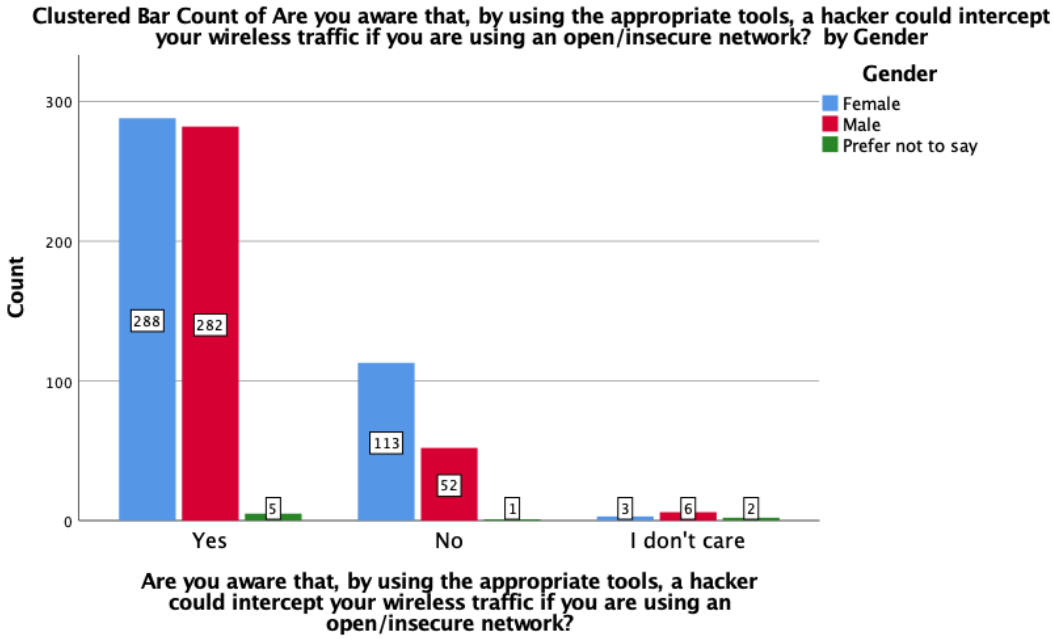


Figure 4-32: Clustered bar chart showing breakdown of respondents' awareness of hacker intercepting traffic over open wireless network by gender

4.11 Data Storage

Respondents were asked if they used a USB key (pen drive) or an external hard drive to store data for the purpose of storing data for their relevant course. 63% (476 respondents out of 752) stated that they used either a USB pen drive or an external hard drive. These 476 respondents were asked with a follow up question whether or not the external drive or USB key they used was encrypted. As can be seen from Figure 4-33 below, only 18.9% of respondents who used one of these devices claimed that the device was encrypted.

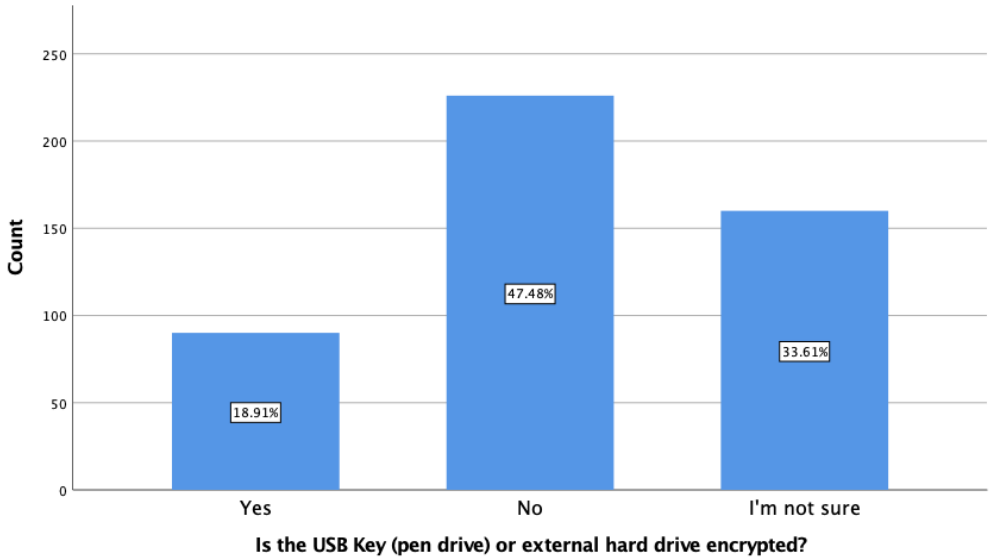


Figure 4-33: Breakdown of respondents who encrypt USB key/hard drive

A clustered bar chart was used to show a further breakdown of these figures by gender. Although the number of male and female respondents who stated that the device was not encrypted was evenly matched, a higher proportion of females said they did not know if the device was encrypted.

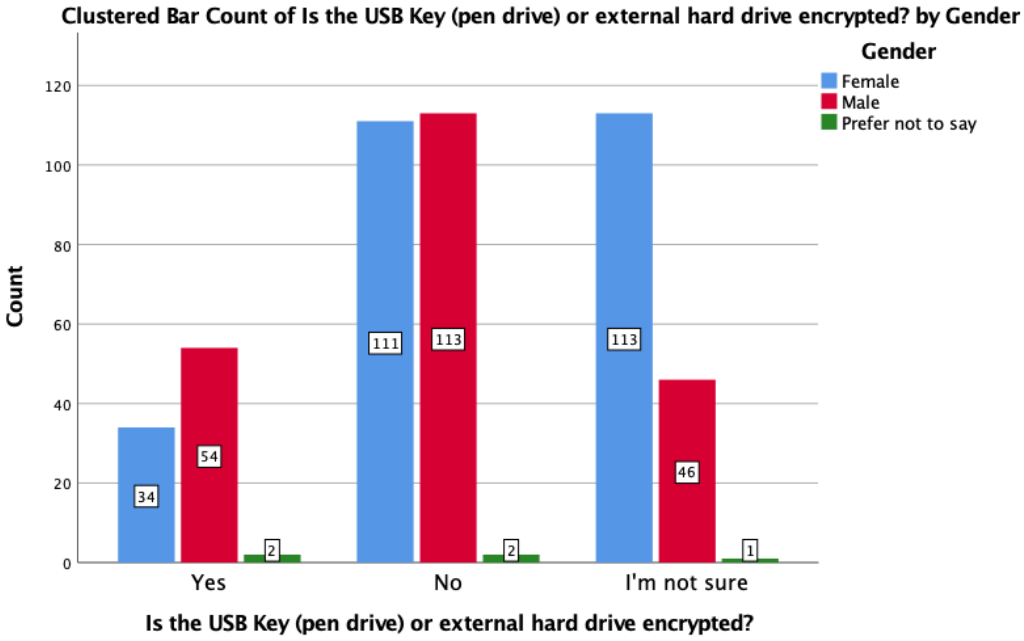


Figure 4-34: Breakdown of respondents who encrypt USB key/hard drive by gender

This data was also broken down by area of study, with a significantly higher number of students within the College of Engineering and Built environment stating that the device was not encrypted, as well as students based in the Science and Health. These figures can be seen in Figure 4-35 below.

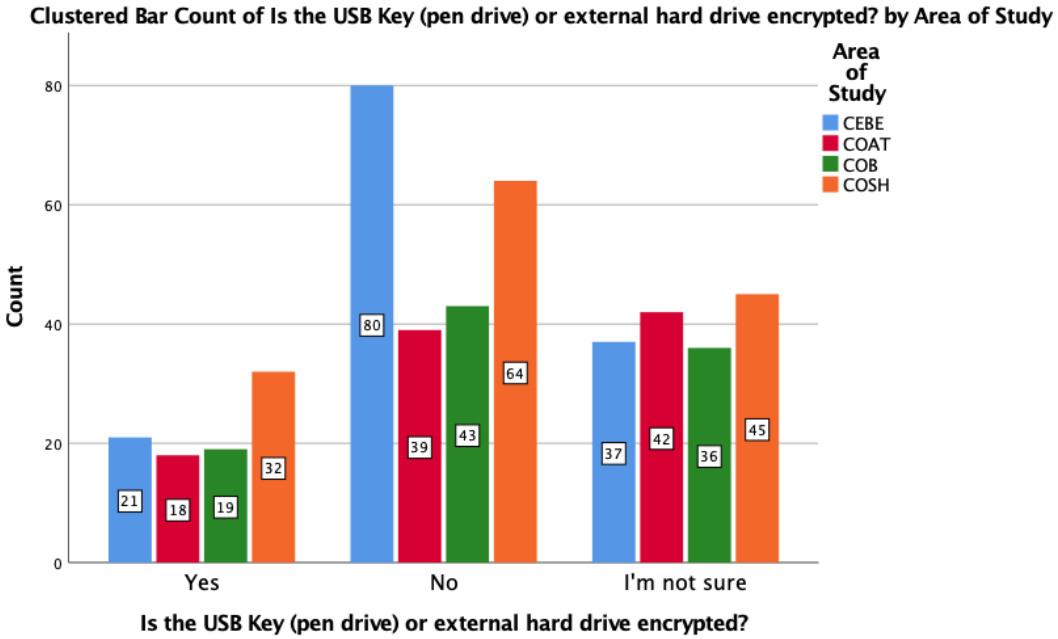


Figure 4-35: Breakdown of respondents who encrypt USB key/hard drive by area of study

4.12 Summary

In this chapter, the respondents quiz scores and behaviours were assessed and presented in a number of graphs and charts. The various demographic breakdown was presented to show the number of respondents for each category. The IT competency of each user was reviewed, along with a summary of users who had previously been involved in a security breach.

Respondent's behaviours were presented in relation to their own device habits, password habits and use and awareness of security features such as password managers and multi-factor authentication. Due to space constraints, not all results were presented in this chapter that were captured in this survey. Additional results showing the demographic breakdown relating to device habits, including awareness of OS updates, software updates and Anti-virus updating can be found in Appendix C.

5 ANALYSIS & EVALUATION

5.1 Introduction

This chapter will discuss and analyse the various results obtained from the survey in order to determine if each of the null hypothesis outlined at the beginning of this document can be either accepted or rejected. The three null hypotheses are listed below:

Hypothesis 1:
H0: When given a quiz relating to IT Security awareness, there will be no significant difference in the mean scores for the various demographic groups

Hypothesis 2:
H0: When respondents' security behaviours and habits are weighted and scored, there will be no significant difference in the mean scores for the various demographic groups

Hypothesis 3:
H0: There will be a significant relationship between users who claim they have a high level of information security awareness and those who have received the actual training

5.2 Hypothesis 1: scenario-based quiz

A total of twelve multiple choice questions were presented at the end of the survey. Each question described a scenario and asked the respondent to select the answer they deemed to be the most appropriate and the most secure in that particular scenario. The answers to each question were set to be in a random order each time the survey was completed. A full list of these behavioural analysis questions asked in the survey can be found in **Appendix A – Survey Questions.**

5.2.1 Summary of quiz scores

Error! Reference source not found.1 below shows how many questions each respondent answered correctly during the behaviour analysis section.

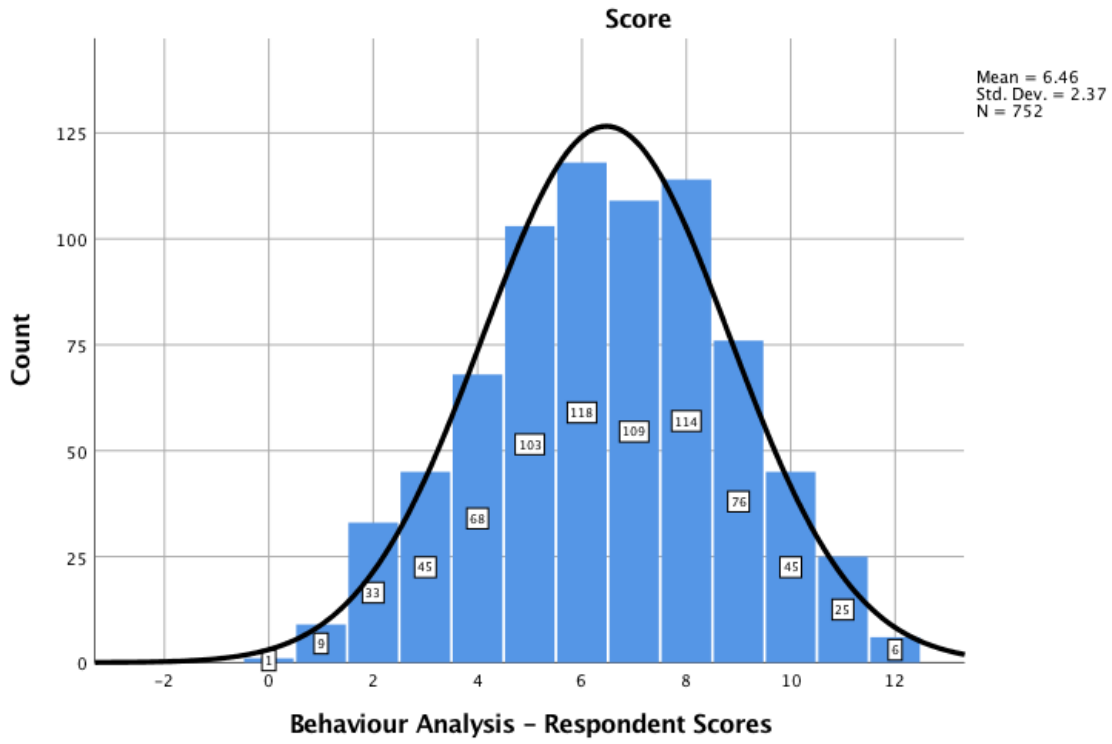


Figure 5-1: Summary of respondents' quiz scores

Only six respondents (0.8%) obtained a perfect score of 100% (12/12), with one respondent managing to score 0% (0/12). The mean score was 6.46, with the median score being 6/12. Figure 5-2 below shows the cumulative distribution of the quiz scores obtained by respondents.

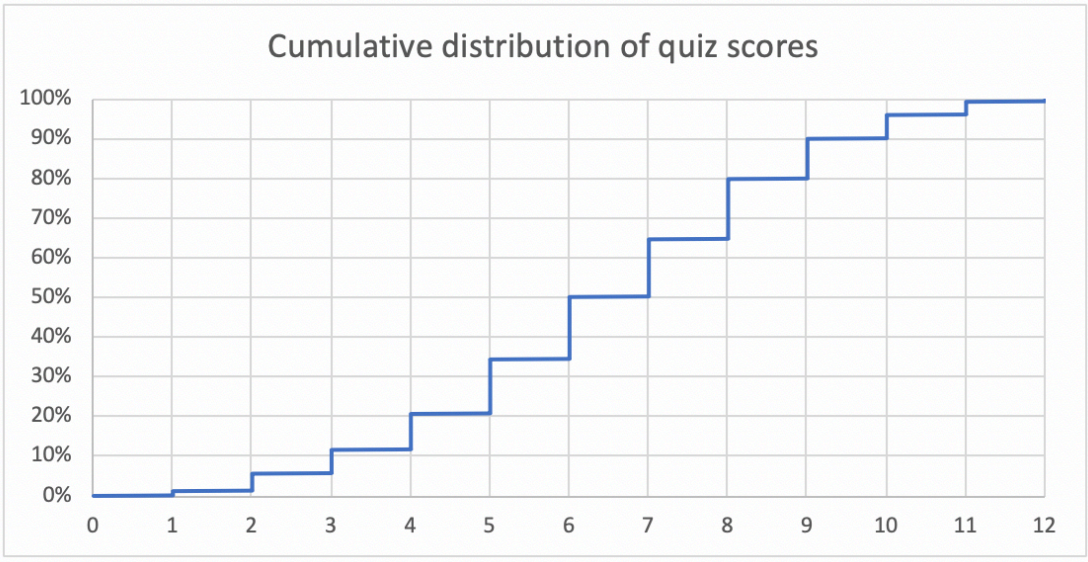


Figure 5-2: Cumulative distribution of scores relating to quiz scores

As outlined in the design and methodology section of this document, the questions asked in the quiz related to four different categories: passwords and MFA; wireless technologies; phishing and email and data protection.

Figure 5-3 below gives a breakdown for how each question was answered. The number of correctly answers questions are highlighted in blue, with the number of incorrect answers highlighted in red. A number of questions also gave an option for the respondent to answer the question with "I don't know". These responses are highlighted in green. Not all questions gave this as an option, only questions 2, 5, 6, 7, 10 and 12 gave the option for the user to state they did not know the answer.

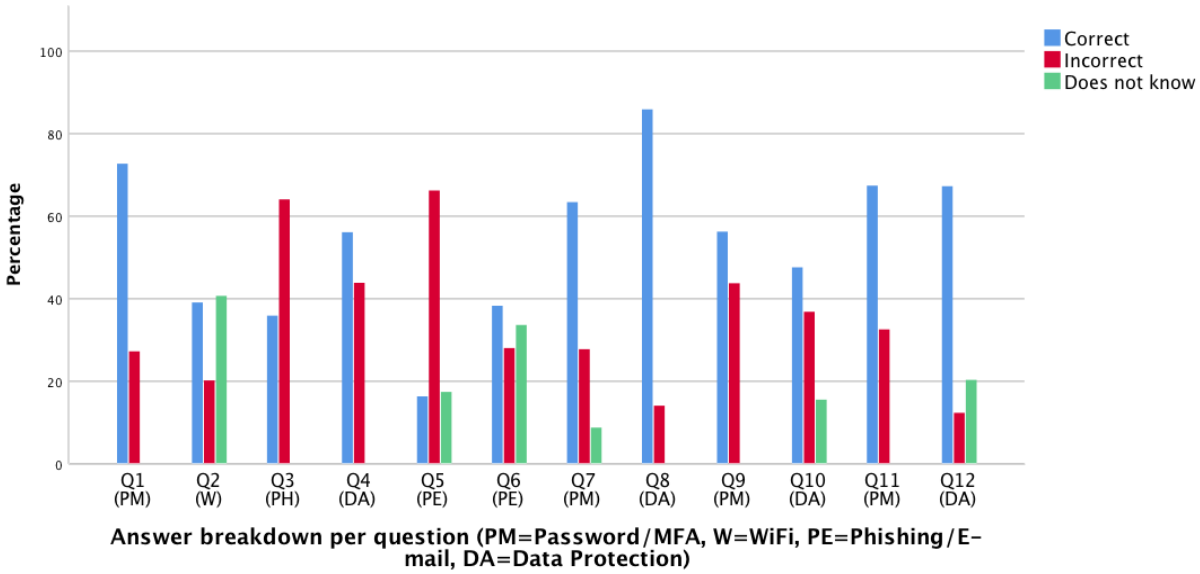


Figure 5-3: Breakdown of correct and incorrect answers per question

As can be seen in Figure 5-3 above, question three and question five had a high number of incorrect answers, with a high number of respondents answering question eight correctly.

A complete listing of these statistics can be seen in Table 5-1 below

	Correct	Incorrect	Don't know
Q1 (PM)	72.70%	27.30%	
Q2 (W)	39.10%	20.20%	40.70%
Q3 (PE)	35.90%	64.10%	
Q4 (DA)	56.10%	43.90%	
Q5 (PE)	16.40%	66.20%	17.40%
Q6 (PE)	38.30%	28.10%	33.60%
Q7 (PM)	63.40%	27.80%	8.80%
Q8 (DA)	85.90%	14.10%	
Q9 (PM)	56.30%	43.80%	
Q10 (DA)	47.60%	36.80%	15.60%
Q11 (PM)	67.40%	32.60%	
Q12 (DA)	67.30%	12.40%	20.30%

(PE) Phishing attempts & email (3 questions) (Q3) (Q5) (Q6) / (W) Wireless technology (1 question) (Q2) / (PM) Passwords/MFA (4 question) (Q1) (Q7) (Q9) (Q11) / (DA) Data Protection (4 question) (Q4) (Q8) (Q10) (Q12)

Table 5-1: breakdown of correct and incorrect answers per question

5.2.2 ISA Self-assessment comparison with mean scores

Before the various demographic groups were compared to determine if there were any significant differences between the mean scores obtained in the quiz, the mean scores were compared with the self-assessment score of each respondent. As can be seen in Table 5-2 below, the mean score increases with the ISA self-assessment rating. This confirms that there is high degree of honesty from respondents when they completed the survey.

ISA – Self Assessment	**Participants**	**Mean Score**	**Std. Deviation**
Very Poor	21	4.14	2.151
Poor	80	5.63	2.046
Fair	182	5.78	2.091
Good	206	6.26	2.213
Very Good	159	7.35	2.309
Excellent	79	7.68	2.222
Exceptional	25	8.24	2.788

Table 5-2: Comparison of ISA self-assessment with mean scores

5.2.3 Demographic analysis - Gender

An independent t-test, also known as a two-sample t-test was used to determine if there was a significant difference of mean scores obtained in the quiz between male and female respondents.

As part of this research, the null hypothesis stated the following:

H0: When given a quiz relating to IT Security awareness, there will be no significant difference in the mean scores for the various demographic groups.

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Score	Male	340	7.12	2.461	.133
	Female	404	5.91	2.134	.106

Table 5-3: Descriptive statistics of quiz scores obtained by respondents by gender

A total of eight respondents did not wish to state their gender as part of the survey. These numbers were removed from the figure below in order to determine the mean score between male and female respondents. As can be seen in Table 5-3 above, the mean score for female respondents was 5.91, whereas male respondents scored a higher mean of 7.12.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Score	Equal variances assumed	8.764	.003	7.228	742	.000	1.218	.168	.887	1.548
	Equal variances not assumed			7.140	676.008	.000	1.218	.171	.883	1.552

Table 5-4: Result of independent t-test for variation of scores by gender for behaviour analysis

Using the formula to determine t-value in an independent t-test, we can determine that the t value = 7.140, the p-value is < .0001, meaning the result is significant at $p < .05$. We can therefore reject the null hypothesis, as there is a significant difference between the scores obtained by male and female respondents.

5.2.4 Demographic analysis - Age

In order to determine if there was a significant difference in the mean quiz scores obtained in the survey amongst the remaining demographic groups, a one-way ANOVA test was carried out on these variables. Respondents were asked to select an age category during the survey to identify their age. Table 5-5 below gives a breakdown of the mean score obtained from each range, along with the standard deviation of each group.

Age Range	Mean	N	Std. Deviation
17-19	6.20	136	2.280
20-21	6.22	209	2.220
22-23	6.49	138	2.429
24-27	6.69	88	2.346
28-34	6.74	58	2.763
35-44	6.89	64	2.344
45-54	7.12	40	2.399
55+	5.82	11	2.750
Total	6.46	744	2.366

Table 5-5: Descriptive statistics of quiz scores obtained by respondents by age range

Running a one-way ANOVA test on these age ranges, we can see the results in Table 5-6 below. The f-ratio value is 1.680. The p-value = 0.111. This means that the result is not significant at $p < .05$.

Score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	65.420	7	9.346	1.680	.111
Within Groups	4095.526	736	5.565		
Total	4160.946	743			

Table 5-6: Result of one-way ANOVA test to determine significance of quiz scores by age group

Due to these results, we can accept the null hypothesis outlined above

5.2.5 Demographic Analysis - Education

The area of study, education level and whether the student was full-time or part-time was also examined as part of this analysis. The first demographic examined in the area of education was to run a comparison between full-time and part-time students.

5.2.5.1 Student status

As there were only 2 values being compared, a two-sample t-test was used to determine if there was a significant difference between the mean scores obtained in the quiz between full-time and part-time students. Table 5-7 shows the mean score for each group.

	Are you studying full time or part time?	N	Mean	Std. Deviation	Std. Error Mean
Score	Full-time	606	6.27	2.316	.094
	Part-time	138	7.30	2.415	.206

Table 5-7: Descriptive statistics of quiz scores obtained by full-time and part-time students

Using a two-sample t-test in SPSS, we can determine that t=-4.534, the p-value is < .0001, meaning the result is significant at $p < .05$. We can therefore reject the null hypothesis, as there is a significant difference between the scores obtained by full-time students and part-time students. These results can be seen below in Table 5-8.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Score	Equal variances assumed	.645	.422	-4.655	742	.000	-1.025	.220	-1.457	-.593
	Equal variances not assumed			-4.534	198.418	.000	-1.025	.226	-1.471	-.579

Table 5-8: Result of independent t-test for variation of scores by full-time and part-time status

5.2.5.2 Area of study

The area of study was also compared to determine if there was a significant difference between the four major disciplines within TU Dublin. Table 5-9 below outlines the means score of students within each discipline.

Area of Study	Mean	N	Std. Deviation
CEBE	6.70	188	2.449
COAT	6.05	139	2.165
COB	6.04	183	2.271
COSH	6.85	234	2.409
Total	6.46	744	2.366

Table 5-9: Descriptive statistics of quiz scores obtained by respondents by Area of study

A one-way ANOVA was carried out on these areas of study. The f-ratio value is 6.185. The p-value = 0.0001. This means that the result is significant at $p < .05$. Due to these results, the null hypothesis cannot be accepted in relation to area of study, as there is a significance the scores obtained between students in the various disciplines.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	101.786	3	33.929	6.185	.000
Within Groups	4059.160	740	5.485		
Total	4160.946	743			

Table 5-10: Result of one-way ANOVA test to determine significance of scores by area of study

5.2.5.3 Level of Study

A similar one-way ANOVA was carried out on the results of the level of study. Table 5-11 below gives a breakdown of the mean scores obtained from each group.

Current Level of Study	Mean	N	Std. Deviation
Apprenticeship / Trades	5.50	2	3.536
Graduate (Masters)	6.66	124	2.650
Post Graduate (PhD)	6.25	20	2.633
Undergraduate (1st Year)	6.14	171	2.260
Undergraduate (Year 2, 3 or 4)	6.55	427	2.301
Total	6.46	744	2.366

Table 5-11: Descriptive statistics of quiz scores obtained by respondents by level of study

A one-way ANOVA test was carried out on these level of study categories. As can be seen in Table 5-12 below, the f-ratio = 1.275, p-value = 0.278. This means that the result is not significant at $p < 0.5$. We can therefore accept the null hypothesis in relation to the level of study.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28.525	4	7.131	1.275	.278
Within Groups	4132.422	739	5.592		
Total	4160.946	743			

Table 5-12: Result of one-way ANOVA test to determine significance of scores by level of study

If we remove the Apprenticeship / Trades from the above results, the result is still regarded as not significant with the f-ratio = 1.592 and p=0.190.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	26.667	3	8.889	1.592	.190
Within Groups	4119.922	738	5.583		
Total	4146.589	741			

Table 5-13: Result of one-way ANOVA test to determine significance of scores by level of study, without apprentices

5.2.6 Summary of results

Demographic	p-value	Significant	Accept null hypothesis
Gender	< .00001	Yes	No
Age	0.111	No	Yes
Student status (FT/PT)	< .0001	Yes	No
Area of Study	< .0001	Yes	No
Level of study	0.278	No	Yes

Table 5-14: Summary of results to determine if results of each test was significant

As can be seen from Table 5-14 above, we can observe that there is a significant difference in the mean scores obtained by male and female students, full-time and part-time students, as well as the area of study each student is involved with when respondents were given a quiz relating to IT Security awareness. There was no significant difference between the various age groups, nor was there a significant difference when the level of study was assessed.

5.3 Hypothesis 2: Behaviour analysis

As part of the survey, each respondent was asked to specify their security habits relating to their own personal devices, as well as their password habits relating to their student account and personal online accounts. This part of the researched examined the following null hypothesis:

H0: When respondents' security behaviours and habits are weighted and scored, there will be no significant difference in the mean scores for the various demographic groups.

5.3.1 Data clean-up

In order to determine if this null hypothesis can be accepted or rejected, the security habits of each respondents was assessed in relation to their device usage habits and their password habits. Not every respondent stated they had a personal device that they used for the purpose of completing college assignments. Students that stated that they did not have a device were excluded from this part of the research.

A total of 15 respondents stated that they did not own a personal device, with a total of 41 respondents claiming that they used something other than a PC laptop or Apple Mac Laptop. When these two groups were removed from the data, the total number of respondents left were 696 (n=696).

As part of this analysis, respondents were also asked questions relating to their password habits. Two of the questions relating to password habits allowed the respondent to answer the question with the response of "I would rather not say". This related to the respondent giving details as to how long their password were, as well as details on if they used the same password on multiple sites. As this information was not disclosed by the respondent, it would not be possible to weight the scores assigned with leaving this information in the analysis. It was, therefore, necessary to remove this data from this part of the analysis. A total of 81 respondents answered "I would rather not say" when asked about their password length, with a total of 54 answering the same way when asked if they had used the same password on multiple sites. This resulted in a total of 590 respondents that were able to be assessed for this part of the analysis.

5.3.2 Chi square test

In order to determine if these 590 respondents were representative of the initial sample of 752 respondents obtained from the survey, a chi-square test was performed for each demographic group. Further details on how this chi square was performed can be found in Appendix B. A summary of these values is presented in Table 5-15 below.

Demographic	**p-value**	**Significant**	**Representative of sample**
Gender	< .05	No	Yes
Age	< .05	No	Yes
Student status (FT/PT)	= 0	No	Yes
Area of Study	< .05	No	Yes
Level of study	< .05	No	Yes

Table 5-15: Summary of P-value obtained from Chi Square test comparing Subset of respondents with that of sample obtained from survey

5.3.3 Summary of behaviour analysis scores

In order to assess these habits, each respondent was awarded a score depending on how they answered the question. A breakdown of the questions assessed for this scoring are outlined in table 3-3 and table 3-4 within the design and methodology chapter of this document. A total of 14 questions in the survey were used to score each respondent based on their security habits, particularly in relation to their device and password habits. A maximum score of 16 was achievable, with a minimum score of -1. A breakdown of these scores is presented in Figure 5-4 below.

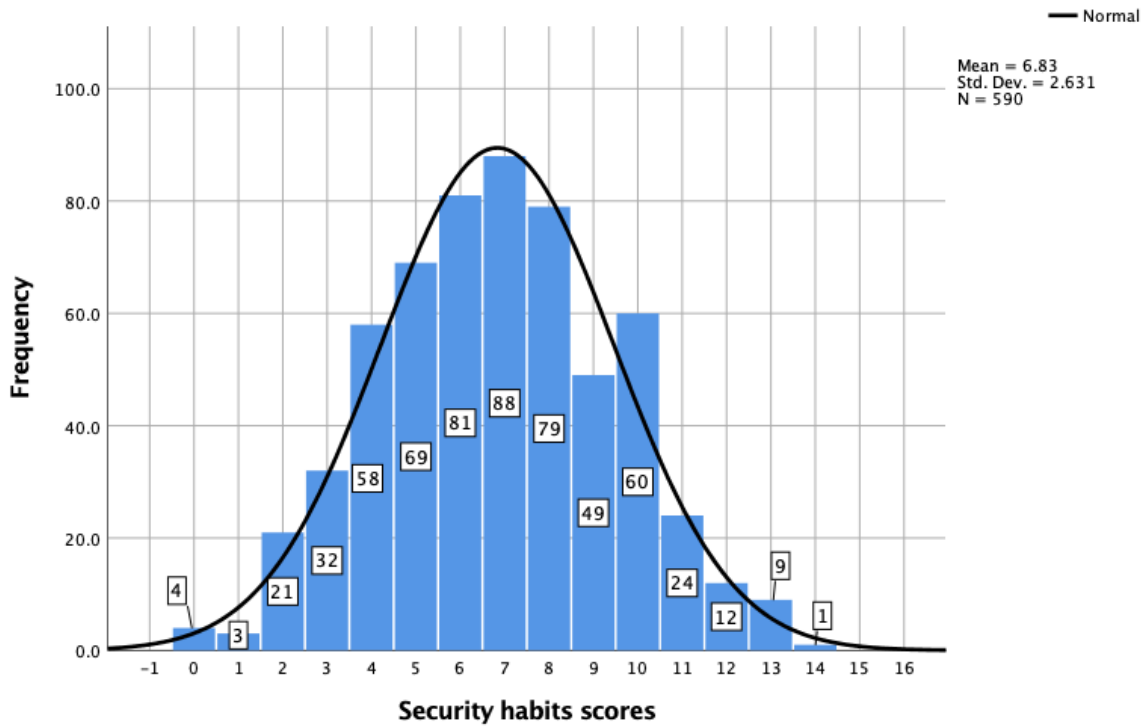


Figure 5-4: Summary of respondents' scores on security habits

As can be seen in Figure 5-4 above, the maximum score obtained was 14. This was obtained by only respondent, with no respondents managing to obtain the maximum score of 16. The lowest score obtained was zero, which was obtained by 4 respondents. The mean score obtained was 6.83 with a standard deviation of 2.63

Figure 5-5 below shows the cumulative distribution of the behavioural scores obtained by respondents.

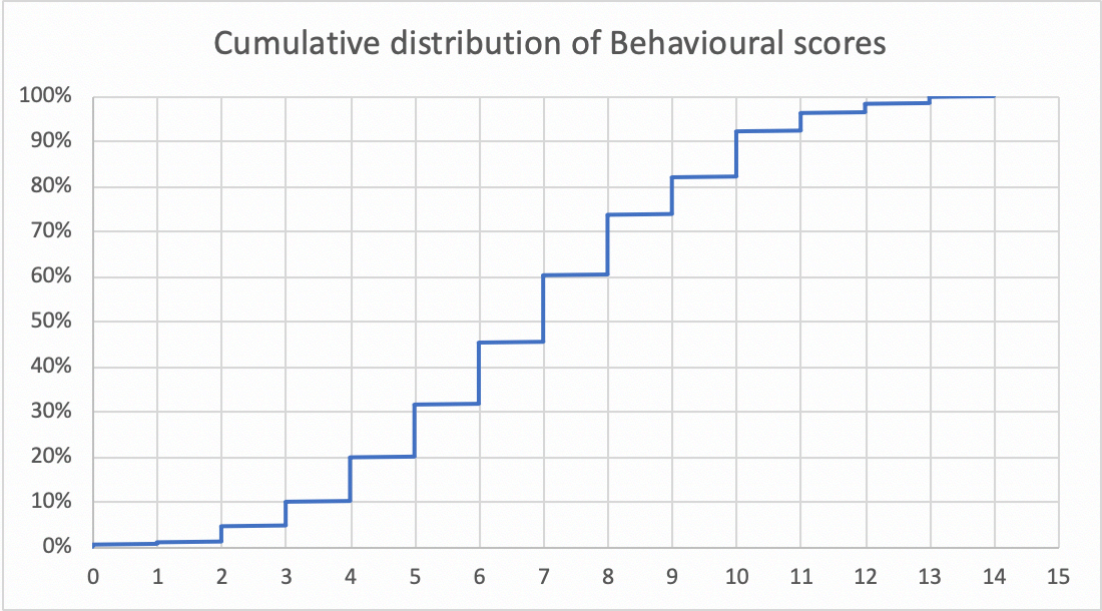


Figure 5-5: Cumulative distribution of scores relating to behaviours

5.3.4 Demographic analysis - Gender

Similar to when the quiz results were analysed, an independent t-test, also known as a two-sample t-test, was used to determine if there was a significant difference between the mean scores of both male and female respondents.

Out of 590 respondents being analysed, a total of seven stated that they did not wish to disclose their gender. When these respondents were removed for this part of the analysis, this gave an overall total of 583 respondents. This number consisted of 336 females and 247 males (n=583)

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Device usage and password habit scores	Male	247	7.61	2.560	.163
	Female	336	6.21	2.511	.137

Table 5-16: Descriptive statistics of behavioural scores obtained by respondents -by gender

As can be seen in Table 5-16 above, female respondents had a mean score of 6.21, with male respondents scoring slightly higher with 7.61.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Device usage and password habit scores	Equal variances assumed	.200	.655	6.583	581	.000	1.397	.212	.980	1.814
	Equal variances not assumed			6.563	524.376	.000	1.397	.213	.979	1.815

Table 5-17: Result of independent t-test for variation of security habit scores by gender

Using the formula to determine t-value in an independent t-test, we can determine that the t value = 6.583, the p-value is < .0001, meaning the result is significant at $p < .05$. We can therefore reject the null hypothesis, as there is a significant difference between the scores obtained by male and female respondents with regard to their security habits.

5.3.5 Demographic analysis – Age

In order to determine if there was a significant difference in the behavioural scores obtained in the survey amongst the remaining demographic groups, a one-way ANOVA test was carried out on these variables. Respondents were asked to select an age category during the survey to identify their age. Table 5-18 below gives a breakdown of the mean score obtained from each age range, along with the standard deviation of each group.

Descriptive Statistics

Age Range		N	Mean	Std. Deviation
17-19	Device usage and password habits scores	116	6.34	2.257
20-21	Device usage and password habits scores	169	6.51	2.598
22-23	Device usage and password habits scores	102	6.68	2.430
24-27	Device usage and password habits scores	67	7.24	2.438
28-34	Device usage and password habits scores	42	7.17	2.853
35-44	Device usage and password habits scores	49	7.82	3.019
45-54	Device usage and password habits scores	35	7.74	3.230
55+	Device usage and password habits scores	10	7.60	3.239

Table 5-18: Descriptive statistics of behavioural scores obtained by respondents by age range

As can be seen in the table above, the age range of 35-44 scored the highest with a mean of 7.82, with the lowest scores being observed in the 17-19 age range, which had a mean score of 6.34.

ANOVA

Device usage and password habit scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	146.829	7	20.976	3.107	.003
Within Groups	3928.893	582	6.751		
Total	4075.722	589			

Table 5-19: Result of one-way ANOVA test to determine significance in behaviour by age group

Running a one-way ANOVA test on these age ranges, we can see the results in Table 5-19 above. The f-ratio value is 3.107. The p-value = 0.003. This means that the result is significant at $p < .05$. We can therefore reject the null hypothesis that there should be no significant difference with regard to security habits between the various age ranges as the result above shows that there is a significant difference amongst the various age ranges.

5.3.6 Demographic analysis – Education

Similar to the how the quiz scores were analysed, the area of study, education level and whether the student was full-time or part-time was also examined as part of this analysis. The first demographic examined in the area of education was to run a comparison between full-time and part-time students.

5.3.6.1 Student status

As there are only two variables being compared, a two-sample t-test was used to determine if there was a significant difference between the mean scores obtained between full time and part time students in relation to their security habits. Table 5-20 below shows the mean score for each group.

Group Statistics

	Status	N	Mean	Std. Deviation	Std. Error Mean
Device usage and password habit scores	Full-time student	479	6.62	2.553	.117
	Part-time student	111	7.77	2.767	.263

Table 5-20: Descriptive statistics of behaviour scores obtained by full-time and part-time students

As can be seen in Table 5-20 above, and similar to the quiz scored analysed in section 5.2, part-time students had a higher mean score compared to full-time students.
Using a two-sample t-test in SPSS, we can determine that t=-3.995 when equal variances are not assumed, the p-value is < .0001, meaning the result is significant at $p < .05$. We can therefore reject the null hypothesis, as there is a significant difference between the security scores obtained by full-time students and part-time students.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Device usage and password habit scores	Equal variances assumed	1.454	.228	-4.200	588	.000	-1.148	.273	-1.685	-.611
	Equal variances not assumed			-3.995	156.291	.000	-1.148	.287	-1.715	-.580

Table 5-21: Result of independent t-test for variation in behaviour scores by full-time and part-time status

5.3.6.2 Area of study

The next section analysed for this part of the analysis was to do with the area of study to determine if there was a significant difference between the four major disciplines within TU Dublin. Table 5-22 below outlines the means score of students within each discipline.

Descriptive Statistics

Area of Study		N	Mean	Std. Deviation
CEBE	Device usage and password habit scores	139	7.06	2.662
COAT	Device usage and password habit scores	115	6.70	2.399
COB	Device usage and password habit scores	148	6.48	2.651
COSH	Device usage and password habit scores	188	7.03	2.710

Table 5-22: Descriptive statistics of behavioural scores obtained by respondents by Area of study

As can be seen in Table 5-22 above, there was a slightly higher mean score obtained by students based on the College of Engineering and Built environment compared to the other three colleges. In comparison, College of Science and Health students averaged a higher mean in relation to the quiz.

A one-way ANOVA was carried out on this data to determine if there was a significant difference in the mean scores. As we can see in Table 5-23 below, the f-ratio value found by this test was = 1.696. The p-value = 0.167. This means that the result is not significant at $p < .05$. Due to these results, we can accept the null hypothesis that there is no significant difference in the mean scores relating to security habits between respondents of the various areas of study.

ANOVA

Device usage and password habit scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	35.087	3	11.696	1.696	.167
Within Groups	4040.635	586	6.895		
Total	4075.722	589			

Table 5-23: Result of one-way ANOVA test to determine significance in behaviour by area of study

5.3.6.3 Level of study

The last demographic group to be analysed is related to the level of study each respondent is currently at. A similar one-way ANOVA was carried out on this data to determine if there was a significant difference between these groups in relation to their security habits. Table 5-24 below gives an overview of the mean scores obtained by each group.

Descriptive Statistics

Current Level of Study		N	Mean	Std. Deviation
Apprenticeship / Trades	Device usage and password habit scores	1	7.00	.
Graduate (Masters)	Device usage and password habit scores	97	7.38	2.395
Post Graduate (PhD)	Device usage and password habit scores	17	6.94	3.325
Undergraduate (1st Year)	Device usage and password habit scores	142	6.75	2.430
Undergraduate (Year 2, 3 or 4)	Device usage and password habit scores	333	6.71	2.733

Table 5-24: Descriptive statistics of behavioural scores obtained by respondents by level of study

As can be seen in the table above, graduate students had a higher mean score compared to the other groups. A one-way ANOVA test was carried out on these categories. As can be seen in Table 5-25 below, the f-ratio = 1.298, p-value = 0.269. This means that the result is not significant at $p < 0.5$. We can therefore accept the null hypothesis that there is no significant difference in the mean scores relating to security habits between respondents of the various level of study.

ANOVA

Device usage and password habit scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	35.862	4	8.965	1.298	.269
Within Groups	4039.860	585	6.906		
Total	4075.722	589			

Table 5-25: Result of one-way ANOVA test to determine significance in behaviour by level of study

Due to only one respondent being within the category of "Apprenticeship / trades", a one-way ANOVA was performed without this group included to confirm if this group was skewing the results. As can be seen below in Table 5-26, p-value obtained with this

group excluded was = 0.160. Even with this group excluded, there is still no significant difference in the mean scores relating to security habits between respondents of the various level of study.

ANOVA

Habits (Device & Password) scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	35.834	3	11.945	1.730	.160
Within Groups	4039.860	585	6.906		
Total	4075.694	588			

Table 5-26: Result of one-way ANOVA test to determine significance of scores by level of study with Apprenticeships / trades excluded

5.3.7 Summary of Security habit analysis

Table 5-27 below gives an overview of the results obtained from each test carried out to determine if there was a significant difference between the various groups amongst each demographic.

Demographic	**p-value**	**Significant**	**Accept Null Hypothesis**
Gender	< .00001	Yes	No
Age	0.003	Yes	No
Student status (FT/PT)	< .0001	Yes	No
Area of Study	0.167	No	Yes
Level of study	0.269	No	Yes

Table 5-27: Overview of results to determine if there is a significant difference between the various demographic groups in relation to behaviour

As can be seen from Table 5-27 above, we can observe that there is a significant difference in the scores obtained by male and female students, the various age ranges and the full-time / part-time status of each student when respondents were assigned weighted scores in relation to the device usage habits and password habits. There was no significant difference between the area of study or the level of study of each.

5.4 Hypothesis 3: Participation in security awareness training

The third part of this research was to establish if the following null hypothesis should be accepted or rejected:

H0: There will be a significant relationship between users who claim they have a high level of information security awareness and those who have received the actual training

As we demonstrated in section 4 of this document, just over 16% of respondents had stated they had participated in information security awareness (ISA) training either within the last two years or longer than two years ago. 9.3% of respondents stated that they were not sure if they had participated in this type of training.

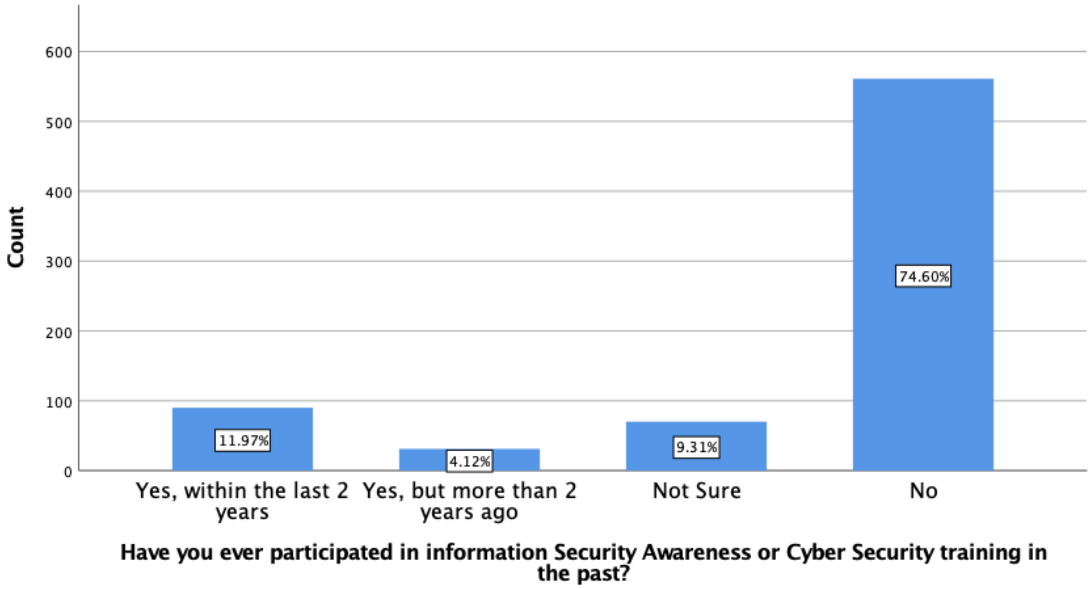


Figure 5-6: Number of respondents that have participated in security awareness training

5.4.1 Comparison of ISA training with quiz scores

Table 5-28 below outlines the mean scores obtained in the quiz between respondents who stated they had participated in information security awareness training in the past.

Have you ever participated in information Security Awareness or Cyber Security training in the past?	Mean	N	Std. Deviation
No	6.25	561	2.366
Not Sure	6.34	70	2.042
Yes, but more than 2 years ago	6.68	31	2.386
Yes, within the last 2 years	7.84	90	2.182
Total	6.46	752	2.370

Table 5-28: Breakdown of respondents' quiz scores in relation to if and when they had participated in information security awareness training

An independent t-test was used to compare respondents that had participated in training, whether it be in the last 2 years or more than 2 years ago, with those who had not participated in any type of training. Table 5-29 below gives the mean scores for respondents that had participated in training with those who did not.

	Participated in information security awareness training	N	Mean	Std. Deviation	Std. Error Mean
Score	Yes	121	7.55	2.284	.208
	No	561	6.25	2.366	.100

Table 5-29: Mean scores obtained by respondents' in relation to those that have and have not participated in training

A total of 70 respondents had answered that that they were not sure if they had participated in any type of training. These respondents were excluded from this analysis.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Score	Equal variances assumed	.890	.346	5.513	680	.000	1.299	.236	.837	1.762
	Equal variances not assumed			5.640	179.895	.000	1.299	.230	.845	1.754

Table 5-30: Comparison of quiz results by participation in training

Table 5-30 above shows us that we can determine that the t value = 5.513, the p-value is < .0001, meaning the result is significant at $p < .01$. This confirms that there is a significant difference between the scores for respondents who have participated in training and those who did not.

5.4.1.1 Comparison of quiz scores of respondents that have participated in training within last 2 years versus more than 2 years ago

A second independent t-test was used to compare respondents that had participated in information security awareness training in the past 2 years with those who had participated in the training more than two years ago. Table 5-31 gives a breakdown of the mean scores obtained by each group.

	If respondent has participated in information security awareness training	N	Mean	Std. Deviation	Std. Error Mean
Score	Participated in training in last 2 years	90	7.84	2.182	.230
	Participated in training more than 2 years ago	31	6.68	2.386	.429

Table 5-31: Comparison of quiz results with respondents that had participated in training within last 2 years compared to more than 2 years ago

As can be seen from Table 5-32 below, the t-value = 2.507, the p-value = 0.02, meaning the result is significant at $p < 0.5$. This confirms that there is also a significant difference between the scores of those who have participated in the training in the past 2 years and those who participated in the training more than 2 years ago.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Score	Equal variances assumed	.545	.462	2.507	119	.014	1.167	.466	.245	2.089
	Equal variances not assumed			2.400	48.424	.020	1.167	.486	.189	2.145

Table 5-32: Comparison of quiz results by participation in training in last 2 years

5.4.2 Comparison of ISA training with security habits

The next part of this analysis examined if there was a significant difference between the mean scores obtained by observing the security habits of each respondent with those that had participated in ISA training with those who have not. Security habits were calculated by weighted scores outlined in Table 3-2 and Table 3-3 on page 53 and 54 of the design and methodology section of this document.

As has been done when assessing this data, respondents who stated that they did not own and use a Windows PC laptop or Apple Mac laptop for college assignments were

excluded from this part of the analysis. Respondents that opted to answer password questions with "would prefer not to say" were also removed, which resulted in a total of 590 respondents being analysed for this part of the research. Table 5-33 below gives a breakdown of these 590 respondents to show how many have participated in information security awareness training with those who have not.

Have you ever participated in information Security Awareness or Cyber Security training in the past?	Mean	N	Std. Deviation
No	6.61	442	2.510
Not Sure	6.15	55	2.542
Yes, but more than 2 years ago	7.12	25	2.403
Yes, within the last 2 years	8.76	68	2.749
Total	6.83	590	2.631

Table 5-33: Breakdown of respondents' security habit scores in relation to if and when they had participated in information security awareness training

An independent t-test was used to compare respondents that had participated in training, whether it be in the last 2 years or more than 2 years ago, with those who had not participated in any type of training. A total of 55 respondents stated they did not know if they had participated in ISA training. These were removed from this part of the analysis. Table 5-34 gives the mean scores for respondents that had participated in training with those who did not.

Group Statistics

	Participated in Training (Yes/No)	N	Mean	Std. Deviation	Std. Error Mean
Device usage and password habit scores	Yes	93	8.32	2.747	.285
	No	442	6.61	2.510	.119

Table 5-34: Mean scores of security habits obtained by respondents' in relation to those that have and have not participated in training

As can be seen in the table above, a total of 93 respondents stated that they had participated in some form of ISA training. The mean score obtained by these respondents was considerably higher (8.32) when compared to those who have not participated in training (6.61)

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Device usage and password habit scores	Equal variances assumed	1.812	.179	5.894	533	.000	1.716	.291	1.144	2.288
	Equal variances not assumed			5.557	126.333	.000	1.716	.309	1.105	2.327

Table 5-35: T-test result showing significance in difference of security habits by training

Table 5-35 above shows us that we can determine that the t value = 5.894, the p-value is < .0001, meaning the result is significant at $p < .01$. This confirms that there is a significant difference between the security habits for respondents who have participated in training and those who did not.

5.4.2.1 Comparison of security habit scores of respondents that have participated in training within last 2 years versus more than 2 years ago

A final independent t-test was used to compare respondents that had participated in information security awareness training in the past 2 years with those who had participated in the training more than two years ago. Table 5-36 below gives a breakdown of the number of respondents that had participated in information security awareness within the past 2 years with those who participated in training more than 2 years ago.

Group Statistics

	Have you ever participated in information Security Awareness or Cyber Security training in the past?	N	Mean	Std. Deviation	Std. Error Mean
Device usage and password habits	Yes, but more than 2 years ago	25	7.12	2.403	.481
	Yes, within the last 2 years	68	8.76	2.749	.333

Table 5-36: Comparison of security habit results with respondents that had participated in training within last 2 years compared to more than 2 years ago

As can be seen from Table 5-37 below, when a t-test was run on this data, the t-value = -2.812, the p-value = 0.007, meaning the result is significant at $p < 0.1$. This confirms that there is also a significant difference between the security habit scores of those who have participated in the training in the past 2 years and those who participated in the training more than 2 years ago.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Device usage and password habits	Equal variances assumed	1.309	.256	-2.642	91	.010	-1.645	.623	-2.881	-.408
	Equal variances not assumed			-2.812	48.606	.007	-1.645	.585	-2.820	-.469

Table 5-37: Comparison of security habit results by participation in training in last 2 years

5.4.3 Summary of results comparing results of security habits with participation in training

Although less than 16% of respondents who took part in the survey stated that they had participated in information security awareness training in the past, the figures show that that there was a significant difference in the mean scores when a comparison was done on both the security habits and the behaviour analysis of respondents. The figures also show that respondents that had participated in this type of training within the past 2 years scored significantly higher than those who had participated in the training more than 2 years ago.

5.5 Summary of Analysis and Evaluation

This part of the research examined if there was a significant difference between certain demographic groups when these respondents were given a quiz relating to security awareness. The results identified that there were significant differences amongst male and female students, as well as full-time and part-time students. Students who study in the area of Science and Health, as well as Engineering and built environment scored higher than students involved in applied arts or business courses.

It also found significant differences between certain demographic groups when their own security habits were analysed. This involved analysing individual habits relating to their own devices, as well as looking at their habits when it came to creating and managing their own passwords. The results show there were significant differences between male and female students, full-time and part-time students, as well as students of a certain age. Students in the age range of 17-19 scored considerably less than any other group when their own security habits were compared.

Overall, Male students scored higher in both the quiz and their own security habits, with part-time students scoring higher in both areas compared to full-time students.

The final part of this analysis compared the scores obtained by students who had participated in ISA training with those who had not. The results highlight significant differences in the mean scores between these two groups. This provides some evidence that students who participate in this type of training have a better awareness of information security, but also implement these best practices on their own device.

6 CONCULSIONS AND FUTURE WORK

6.1 Introduction

This chapter of the research will present an overview of the findings, along with the limitations of this research. It will also look at what contributions were made to the body of knowledge and briefly look at what future work could be undertaken in this area.

6.2 Research Overview

As discussed in the literature review of this document, a number of different threats exist which allow cybercriminals to either steal data or gain unauthorised access to a system. The weakest link in any organisation is the end user in terms of computer security countermeasures (Rhodes, 2001). In order to reduce these risks, it is imperative that users are aware of these risks as well as security best practices. The objective of this research was to investigate the following research question:

Are there certain demographic groups within a third level educational institute that have a lower level of information security awareness?

Three separate hypotheses were identified as part of this study. The first was to establish if there was a difference amongst demographic groups when their security behaviours were analysed and weighted. The second was to establish if there was a difference amongst demographic groups when they were quizzed on certain scenarios related to security awareness best practices. The third hypothesis was to establish if there was any relationship between respondents that had undertaken information security awareness training and their own security habits

Quantitative analysis was carried out on the data gathered from the survey in order to determine if these three hypotheses could be accepted or rejected. A number of statistical methods were used to assess these values in order to determine if there were significant differences amongst these groups.

6.3 Limitations of Research

The survey was distributed to all students within TU Dublin (city centre campus) via email. Ideally, some form of random sampling would have been used instead of convenience sampling in order to obtain a better representative of the population in terms of age, gender and area of study. Due to time constraints and the cost involved with implementing this sampling method, the decision was made to go with convenience sampling. Using a random sampling method may have increased the accuracy of the results obtained. Only students were assessed as part of this research. Staff members were not targeted with the survey. There is a need to assess staff amongst a university, as successful phishing attempts on staff working in a financial section of the university could have dire consequences.

Although students were asked if they had participated in ISA training in the past, due to the fact first year undergraduates and part-time students would be part-taking in the survey and there is not a de-facto standard for this type of training, it would have been difficult to ascertain where the student had completed the training as well as the quality of the training. Due to this, it was only possibly to ask if the student had participated in the training or not.

By using a survey to obtain the behavioural analysis and to perform a quiz on each respondent does have some limitations. Firstly, the questions used in the survey needed to be phrased to suit all candidates with a varying degree of IT competency. This meant that there may have been a lack of understanding with some of the terminology used on some of the questions. There is also a number of limitations when using multiple choice questions to assess the level of ISA from candidates. There is a possibility that some respondents would guess an answer correctly without actually knowing it. Although some of the questions gave the option of "I honestly don't know" as a choice for an answer, not all questions listed this, meaning respondents could have accidently selected the correct answer. The use of face to face interviews with students would have allowed for a better understanding of their level of IT competency and eliminated the need to provide multiple choice answers that could be correctly guessed.

Finally, as outlined earlier in this document, the student helpdesk advised that the majority of queries related to accessing services within the University from mobile devices (such as email and access to Wi-Fi). Assessment of mobile device security was not assessed as part of this research, as it was deemed unlikely that students would use these types of devices to complete assignments. Ideally, if time permitted, both mobile and laptop devices would have been assessed as part of this research to give a better overview of the student's security awareness.

6.4 Contributions to the body of knowledge

Previous studies in the area of assessing information security awareness have shown a varying degree of results when experts and non-expert computer users were compared. The purpose of this research was to identify if there were significant differences between certain demographic groups when it came to their own risk behaviours and knowledge on best security practices. The results did highlight that when behaviours were analysed, demographic groups of gender, age and student status were found to be significantly different. It also highlighted that the majority of students within a third level institute do not have the necessary skills or awareness to keep their devices, accounts and data secure. When the quiz scores were compared amongst the various demographic groups, it showed that gender was once again a significant factor, along with the area of study and whether the student was part-time or full-time. Surprisingly, part-time students scored higher in relation to their behaviours and when assessed using the quiz compared to full-time students.

An interesting observation in the survey was that less than 24% of students used a password manager for storing passwords for their online accounts. The majority of students claimed to re-use passwords across different online platforms either all of the time or some of the time.

When reviewing the quiz results, only 16.4% of students were correctly able to validate a legitimate email compared to a phishing email. If the number of students that have never changed their password (49%) is taken into account, this shows that the chances of students being phished for information such as their password is extremely high. As well as this, TU Dublin needs to implement better password policies, and possibly look

to implement MFA on all accounts. Research carried out by Doerfler et al. (2019) shows that by simply adding a recovery phone number to a Google account can block up to 99% of bulk phishing attacks.

6.5 Future Work and Recommendations

This research has primarily focused on student's security habits relating to their own personal device, as well as their password habits. It also assessed their security awareness when quizzed on specific scenarios in relation to security best practices. There were a number of areas related to security in the literature that were not included in the scope of this research. Future work could include areas such as the ability to identify social engineering attacks, identifying risky e-mail attachments and other security aspects related to their mobile phones. It would also be a recommendation to assess the information security awareness of both academic staff and non-academic staff within a third level institute.

Although the findings in this research indicate that there are significant differences between a number of demographic groups, more research is needed to assess the type of training that users are receiving in this area, with a way to quantify if this training is affective on the users attitudes towards their own security habits. Overall, females scored lower than male respondents when their mean scores were compared in relation to their device and password habits and their knowledge on security best practices, but there is little evidence to understand why this is.

The sample size obtained from the survey was relatively high compared to other studies examined in the literature, but it may be more useful to survey all returning students at the start of the next academic year. Future work should look at establishing a customised training module for each demographic group and then re-assess these groups after the training has been provided to verify if there is any improvement in the overall security awareness of students.

With the increased use of cloud services by third level institutes, it may be worthwhile investigating the security risks being taken by the IT departments and decision makers

in these institutes to determine if security best practices are being implemented, and what risks are being taken with offloading student data to third party companies.

6.6 Final thoughts

It may take a significant data breach or some form of financial penalty for third level institutes to start improving security awareness to their student population and to make this type of training mandatory. Both students and staff need to be made aware of the various risks associated with bad practices when it comes to device management and password hygiene. The use and reliance on information technology will continue to grow, as will the number of threats and vulnerabilities. Parallel with these developments, continued research will be necessary to determine if end users have the knowledge and awareness to reduce these risks.

BIBLIOGRAPHY

Afreen, R. (2014). Bring your own device (BYOD) in higher education: opportunities and challenges. *International Journal of Emerging Trends & Technology in Computer Science*, *3*(1), 233–236.

Albrecht, J. P. (2016). How the GDPR will change the world. *Eur. Data Prot. L. Rev.*, *2*, 287.

Albrechtsen, E. (2007). A qualitative study of users' view on information security. *Computers & Security*, *26*(4), 276–289. https://doi.org/10.1016/j.cose.2006.11.004

Albrechtsen, E., & Hovden, J. (2010). Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study. *Computers & Security*, *29*(4), 432–445. https://doi.org/10.1016/j.cose.2009.12.005

Aloul, F. A. (2012). The need for effective information security awareness. *Journal of Advances in Information Technology*, *3*(3), 176–183.

Andrews, D., Nonnecke, B., & Preece, J. (2007). *Conducting Research on the Internet:: Online Survey Design, Development and Implementation Guidelines*. Retrieved from https://auspace.athabascau.ca/handle/2149/1336

Ary, D., Jacobs, L. C., & Razavieh, A. (1996). *Introduction to research in education.. Ft. Worth*. Holt, Rinehart, and Winston, Inc.

Aytes, K., & Connolly, T. (2004). Computer security and risky computing practices: A rational choice perspective. *Journal of Organizational and End User Computing (JOEUC)*, *16*(3), 22–40.

B. Kim, E. (2014). Recommendations for information security awareness training for college students. *Information Management & Computer Security*, *22*(1), 115–126.

Bartlett, J. E., & Ik, J. W. K. (2001). Organizational Research: Determining Appropriate Sample Size. *In Survey Research Information Technology, Learning, and Performance Journal*.

Bosnjak, M., & Tuten, T. L. (2001). Classifying Response Behaviors in Web-based Surveys. *Journal of Computer-Mediated Communication*, *6*(3). https://doi.org/10.1111/j.1083-6101.2001.tb00124.x

Bowden, A., Fox-Rushby, J. A., Nyandieka, L., & Wanjau, J. (2002). Methods for pre-testing and piloting survey questions: illustrations from the KENQOL survey of health-related quality of life. *Health Policy and Planning*, *17*(3), 322–330. https://doi.org/10.1093/heapol/17.3.322

Bulgurcu, B., Cavusoglu, H., & Benbasat, I. (2010). Information Security Policy Compliance: An Empirical Study of Rationality-based Beliefs and Information Security Awareness. *MIS Q.*, *34*(3), 523–548.

Casey, E., Fellows, G., Geiger, M., & Stellatos, G. (2011). The growing impact of full disk encryption on digital forensics. *Digital Investigation*, *8*(2), 129–134. https://doi.org/10.1016/j.diin.2011.09.005

Chen, Q., & 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)*, 454–460. https://doi.org/10.1109/ICMLA.2017.0-119

Choi, B. C. K., & Pak, A. W. P. (2004). A Catalog of Biases in Questionnaires. *Preventing Chronic Disease*, *2*(1). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1323316/

Cochran, W. G. (1977). *Sampling techniques (3rd edition)*. New York: John Wiley & Sons.

Collier, R. (2017). NHS ransomware attack spreads worldwide. *CMAJ*, *189*(22), E786–E787. https://doi.org/10.1503/cmaj.1095434

Cone, B. D., Irvine, C. E., Thompson, M. F., & Nguyen, T. D. (2007). A video game for cyber security training and awareness. *Computers & Security*, *26*(1), 63–72.

Das, A., Bonneau, J., Caesar, M., Borisov, N., & Wang, X. (2014). *The Tangled Web of Password Reuse. 14*, 23–26.

Dillman, D. A. (2011). *Mail and Internet surveys: The tailored design method--2007 Update with new Internet, visual, and mixed-mode guide*. John Wiley & Sons.

Dodge, R. C., Carver, C., & Ferguson, A. J. (2007). Phishing for user security awareness. *Computers & Security*, *26*(1), 73–80. https://doi.org/10.1016/j.cose.2006.10.009

Doerfler, P., Marincenko, M., Ranieri, J., Jiang, Y., Moscicki, A., McCoy, D., & Thomas, K. (2019). *Evaluating Login Challenges as a Defense Against Account Takeover*.

Drevin, L., Kruger, H. A., & Steyn, T. (2007). Value-focused assessment of ICT security awareness in an academic environment. *Computers & Security*, *26*(1), 36–43.

Ehrenfeld, J. M. (2017). WannaCry, Cybersecurity and Health Information Technology: A Time to Act. *Journal of Medical Systems*, *41*(7), 104. https://doi.org/10.1007/s10916-017-0752-1

Florencio, D., & Herley, C. (2007). *A large-scale study of web password habits*. 657–666. ACM.

Fraley, R. C. (2004). *How to conduct behavioral research over the Internet: A beginner's guide to HTML and CGI/Perl*. Guilford Press New York.

Frick, A., Bächtiger, M.-T., & Reips, U.-D. (1999). Financial incentives, personal information and drop-out rate in online studies. *Dimensions of Internet Science*, 209–219.

Furnell, S. M., Bryant, P., & Phippen, A. D. (2007). Assessing the security perceptions of personal Internet users. *Computers & Security*, *26*(5), 410–417. https://doi.org/10.1016/j.cose.2007.03.001

Galesic, M. (2006). Dropouts on the Web: Effects of Interest and Burden Experienced During an Online Survey. *Journal of Official Statistics*, *22*, 313–328.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, *59*(2), 93.

Grassi, P. A., Fenton, J. L., Newton, E. M., Perlner, R. A., Regenscheid, A. R., Burr, W. E., … Choong, Y.-Y. (2017). *NIST Special Publication 800-63b: Digital Identity Guidelines*. Technical report, National Institute of Standards and Technology NIST.

Guo, K. H. (2013). Security-related behavior in using information systems in the workplace: A review and synthesis. *Computers & Security*, *32*, 242–251. https://doi.org/10.1016/j.cose.2012.10.003

Hanus, B., Windsor, J. C., & Wu, Y. (2018). Definition and Multidimensionality of Security Awareness: Close Encounters of the Second Order. *SIGMIS Database*, *49*(SI), 103–133. https://doi.org/10.1145/3210530.3210538

Herath, T., & Rao, H. R. (2009). Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness. *Decision Support Systems*, *47*(2), 154–165. https://doi.org/10.1016/j.dss.2009.02.005

Ion, I., Reeder, R., & Consolvo, S. (2015). *'... No one Can Hack My Mind': Comparing Expert and Non-Expert Security Practices. 15*, 1–20.

Irvine, C. E., & Chin, S.-K. (1998). Integrating security into the curriculum. *Computer*, *31*(12), 25–30.

Johnson, E. C. (2006). Security awareness: switch to a better programme. *Network Security*, *2006*(2), 15–18. https://doi.org/10.1016/S1353-4858(06)70337-3

Joinson, A. N., Reips, U.-D., Buchanan, T., & Schofield, C. B. P. (2010). Privacy, trust, and self-disclosure online. *Human-Computer Interaction*, *25*(1), 1–24. https://doi.org/10.1080/07370020903586662

Jones, B. H., & Heinrichs, L. R. (2012). Do Business Students Practice Smartphone Security? *Journal of Computer Information Systems*, *53*(2), 22–30. https://doi.org/10.1080/08874417.2012.11645611

Joyce, R. A., Powers, J., & Adelstein, F. (2008). MEGA: A tool for Mac OS X operating system and application forensics. *Digital Investigation*, *5*, S83–S90. https://doi.org/10.1016/j.diin.2008.05.011

Katz, F. H. (2005). The effect of a university information security survey on instruction methods in information security. *Proceedings of the 2nd Annual Conference on Information Security Curriculum Development - InfoSecCD '05*, 43. https://doi.org/10.1145/1107622.1107633

Kelley, K., Clark, B., Brown, V., & Sitzia, J. (2003). Good practice in the conduct and reporting of survey research. *International Journal for Quality in Health Care*, *15*(3), 261–266. https://doi.org/10.1093/intqhc/mzg031

Kharraz, A., Robertson, W., Balzarotti, D., Bilge, L., & Kirda, E. (2015). Cutting the Gordian Knot: A Look Under the Hood of Ransomware Attacks. In M. Almgren, V. Gulisano, & F. Maggi (Eds.), *Detection of Intrusions and Malware, and Vulnerability Assessment* (pp. 3–24). Springer International Publishing.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement*, *30*(3), 607–610.

Kruger, H. A., & Kearney, W. D. (2006). A prototype for assessing information security awareness. *Computers & Security*, *25*(4), 289–296. https://doi.org/10.1016/j.cose.2006.02.008

Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F., & Hong, J. (2010). Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology*, *10*(2), 1–31. https://doi.org/10.1145/1754393.1754396

Lalonde Levesque, F., Nsiempba, J., Fernandez, J. M., Chiasson, S., & Somayaji, A. (2013). A Clinical Study of Risk Factors Related to Malware Infections. *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, 97–108. https://doi.org/10.1145/2508859.2516747

Luker, M., & Petersen, R. (2003). *Computer and Network Security in Higher Education*. Retrieved from https://library.educause.edu/resources/2003/1/computer-and-network-security-in-higher-education

Matsuo, H., Mcintyre, K. P., Tomazic, T., & Katz, B. (2004). *The Online Survey: Its Contributions and Potential Problems*.

Maydeu-Olivares, A., & Garcia-Forero, C. (2010). *Goodness-of-Fit Testing*. *7*, 190–196.

McCall, B. (2018). What does the GDPR mean for the medical community? *The Lancet*, *391*(10127), 1249–1250. https://doi.org/10.1016/S0140-6736(18)30739-6

McCarney, D., Barrera, D., Clark, J., Chiasson, S., & van Oorschot, P. C. (2012). Tapas: Design, Implementation, and Usability Evaluation of a Password Manager. *Proceedings of the 28th Annual Computer Security Applications Conference*, 89–98. https://doi.org/10.1145/2420950.2420964

McCormac, A., Zwaans, T., Parsons, K., Calic, D., Butavicius, M., & Pattinson, M. (2017). Individual differences and information security awareness. *Computers in Human Behavior*, *69*, 151–156.

Mitnick, K. D., & Simon, W. L. (2011). *The art of deception: Controlling the human element of security*. John Wiley & Sons.

Moran, P. (2018). *Shining a light on fraud Irish Economic Crime Survey 2018*. Retrieved from Pwc website: https://www.pwc.ie/publications/2018/economic-crime-survey-2018.pdf

Moser, C. A., & Kalton, G. (2017). *Survey methods in social investigation*. Routledge.

Ng, B.-Y., Kankanhalli, A., & Xu, Y. C. (2009). Studying users' computer security behavior: A health belief perspective. *Decision Support Systems*, *46*(4), 815–825.

Öğütçü, G., Testik, Ö. M., & Chouseinoglou, O. (2016). Analysis of personal information security behavior and awareness. *Computers & Security*, *56*, 83–93. https://doi.org/10.1016/j.cose.2015.10.002

Peltier, T. (2000). How to build a comprehensive security awareness program. *COMPUT SECUR J*, *16*(2), 23–32.

Piazza, P. (2006). Security goes to school. *Security Management*, *50*(12), 46–51.

Rezgui, Y., & Marks, A. (2008). Information security awareness in higher education: An exploratory study. *Computers & Security*, *27*(7), 241–253. https://doi.org/10.1016/j.cose.2008.07.008

Rhee, H.-S., Ryu, Y. U., & Kim, C.-T. (2012). Unrealistic optimism on information security management. *Computers & Security*, *31*(2), 221–232.

Rhodes, K. (2001). Operations security awareness: the mind has no firewall. *Computer Security Journal*, *17*(3), 1–12.

Safa, N. S., Solms, R. von, & Futcher, L. (2016). Human aspects of information security in organisations. *Computer Fraud & Security*, *2016*(2), 15–18. https://doi.org/10.1016/S1361-3723(16)30017-3

Schultz, E. (2004). Security training and awareness-fitting a square peg in a round hole. *Computers & Security*, *1*(23), 1–2.

Senthilkumar, K., & Easwaramoorthy, S. (2017a). *A Survey on Cyber Security awareness among college students in Tamil Nadu*. *263*, 042043. IOP Publishing.

Senthilkumar, K., & Easwaramoorthy, S. (2017b). A Survey on Cyber Security awareness among college students in Tamil Nadu. *IOP Conference Series: Materials Science and Engineering*, *263*, 042043. https://doi.org/10.1088/1757-899X/263/4/042043

Sharf, E. (2016). Information exchanges: regulatory changes to the cyber-security industry after Brexit: Making security awareness training work. *Computer Fraud & Security*, *2016*(7), 9–12. https://doi.org/10.1016/S1361-3723(16)30052-5

Sheng, S., Holbrook, M., Kumaraguru, P., Cranor, L. F., & Downs, J. (2010). *Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions*. 373–382. ACM.

Singh, A. S., & Masuku, M. B. (2014). Sampling techniques & determination of sample size in applied statistics research: An overview. *International Journal of Economics, Commerce and Management*, *2*(11), 1–22.

Singh, S. K. (2015). Advantages and Disadvantages of Probability Sampling Methods in Social Research. *National Conference on Innovative Research in Chemical, Physical, Mathematical Sciences, Applied Statistics and Environmental Dynamics*, (CPMSED-2015), 14–18.

Siponen, M. T. (2001). Five dimensions of information security awareness. *SIGCAS Computers and Society*, *31*(2), 24–29.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods* (8th Edition). Ames, Iowa: Iowa State University Press.

Stobert, E., & Biddle, R. (2014). *The password life cycle: user behaviour in managing passwords*. Presented at the Proc. SOUPS.

Tankard, C. (2016). What the GDPR means for businesses. *Network Security*, *2016*(6), 5–8. https://doi.org/10.1016/S1353-4858(16)30056-3

Thomson, M. E., & von Solms, R. (1998). Information security awareness: educating your users effectively. *Information Management & Computer Security*, *6*(4), 167–173.

Todd, M. A., & Guitian, C. (1989). *Computer Security Training Guidelines* (No. NIST Special Publication (SP) 500-172 (Withdrawn); p. 8). https://doi.org/10.6028/NIST.SP.500-172

TU Dublin. (2019, April). TU Dublin News. Retrieved 24 April 2019, from https://www.dit.ie/newsandevents/news/archive2019/news/title168257en.html

Tuten, T. L., Bosnjak, M., & Bandilla, W. (1999). Banner-advertised Web surveys. *Marketing Research*, *11*(4), 16.

van Kessel, P. (2018, October 10). Is cybersecurity about more than protection? Retrieved 9 March 2019, from https://www.ey.com/en_gl/advisory/global-information-security-survey-2018-2019

Vaniea, K. E., Rader, E., & Wash, R. (2014). *Betrayed by updates: how negative experiences affect future security*. 2671–2674. ACM.

Wash, R., Rader, E., Berman, R., & Wellmer, Z. (2016). *Understanding password choices: How frequently entered passwords are re-used across websites*. 175–188.

Whitman, M. E., & Mattord, H. J. (2011). *Principles of information security*. Cengage Learning.

Wilson, M., & Hash, J. (2003). Building an information technology security awareness and training program. *NIST Special Publication*, *800*(50), 1–39.

Woods, N., & Siponen, M. (2019). Improving password memorability, while not inconveniencing the user. *International Journal of Human-Computer Studies*, *128*, 61–71. https://doi.org/10.1016/j.ijhcs.2019.02.003

Zerlang, J. (2017). GDPR: a milestone in convergence for cyber-security and compliance. *Network Security*, *2017*(6), 8–11. https://doi.org/10.1016/S1353-4858(17)30060-0

APPENDIX A

This part of the document contains a list of questions asked in the survey.

Section 1 - Demographics

Gender *

- Male
- Female
- Prefer not to say
- Other...

Age Range *

- 17-19
- 20-21
- 22-23
- 24-27
- 28-34
- 35-44
- 45-54
- 55+

Current Level of Study *

- Undergraduate (1st Year)
- Undergraduate (Year 2, 3 or 4)
- Graduate (Masters)
- Post Graduate (PhD)
- Apprenticeship / Trades

Are you studying full time or part time? *

- ○ Full-time Student
- ○ Part-time Student

Please select one of the following disciplines which best matches your area of study. *

- ○ Applied Arts
- ○ Business
- ○ Engineering
- ○ Science & health
- ○ Other...

Have you ever participated in information Security Awareness or Cyber Security training in the past? *

- ○ Yes, within the last 2 years
- ○ Yes, but more than 2 years ago
- ○ No
- ○ Not Sure

Are you aware that online Security training is available for you to take in TUDublin? *

- ○ Yes
- ○ No

On a scale from 1-7, how do you rate your level of IT competency (1=Very Poor, 2=Poor, 3=Fair, 4=Good, 5=Very Good, 6=Excellent, 7=Exceptional) *

	1	2	3	4	5	6	7	
Very Poor	○	○	○	○	○	○	○	Exceptional

On a scale from 1-7, how do you rate your IT security awareness (1=Very Poor, 2=Poor, 3=Fair, 4=Good, 5=Very Good, 6=Excellent, 7=Exceptional) *

	1	2	3	4	5	6	7	
Very Poor	○	○	○	○	○	○	○	Exceptional

Have you ever experienced a security breach? (e.g. had your email account, online shopping account, online banking account or one of your social media accounts compromised?)

- ○ Yes
- ○ No

Section 2 – Device Usage

Which of the following devices do you own and use as your primary device for college assignments? *

- ○ PC Laptop (Running a window OS)
- ○ Apple Mac Laptop
- ○ I use something else (such as a tablet or a laptop with a different operating System)
- ○ I don't own my own device

PC Laptop

gather details relating to their Windows Laptop

Do you know what version of Windows the Operating System is running on your device? *

- ○ Windows 10
- ○ Windows 8 or 8.1
- ○ Windows 7
- ○ A different version of windows than what is listed above
- ○ I'm not sure

Is your Windows Laptop password protected? (Does it require a windows password to be entered in order to gain access to the desktop) *

- ○ Yes
- ○ No
- ○ I'm not sure

Is your windows laptop encrypted? *

- ○ Yes
- ○ No
- ○ I'm not sure

Do you have Antivirus installed on the device? *

- ○ Yes, I pay for yearly subscription for Antivirus
- ○ Yes, I paid a one off cost to use one
- ○ Yes, I use a free version or one built into the operating system
- ○ No
- ○ I'm not sure

Mac Laptop

Description (optional)

Do you know what version of Mac OS you are running on your Mac Laptop *

- ○ Mac OS 10.14.x
- ○ Mac OS 10.13.x
- ○ Mac OS 10.12.x
- ○ Mac OS 10.11.x
- ○ Mac OS 10.10.x
- ○ A version older than Mac OS 10.10.x
- ○ I'm not sure

Is your Mac Laptop password protected? (Does it require a Username and/or password to be entered in order to gain access to the desktop) *

- ○ Yes
- ○ No
- ○ I'm not sure

Is your Mac OS device encrypted using FileVault? *

- ◯ Yes
- ◯ No
- ◯ I'm not sure
- ◯ I've never heard of FileVault

Do you have Antivirus installed on the device? *

- ◯ Yes, I pay for yearly subscription for Antivirus
- ◯ Yes, I paid a one off cost to use one
- ◯ Yes, I use a free one
- ◯ No, Mac OS devices don't need Anti-virus!
- ◯ I'm not sure

AntiVirus /Anti-Malware

Description (optional)

How often do you update the anti-virus software on your computer (this relates to the anti-virus definitions) *

- ◯ I update them at least once per week
- ◯ I update them when I remember to
- ◯ The Anti-virus software keeps itself up to date
- ◯ I wasn't aware I needed to update them
- ◯ I don't update the Anti-virus software
- ◯ I'm not sure

Do you regularly scan your computer's hard drive for viruses? *

- ◯ Yes, at least once every couple of weeks
- ◯ Yes, the AV software prompts me to do it
- ◯ Yes, the AV software does it itself in the background
- ◯ No, I have never ran a scan of my hard drive
- ◯ I'm not sure

Personal Device (continued)

Description (optional)

How often do you install OS updates? (This includes security updates) *

- ◯ I install them as soon as I am prompted
- ◯ My machine is set to automatically update itself
- ◯ I don't install updates
- ◯ Not sure
- ◯ I don't know what an OS update is

How often do you install software updates on your device (e.g. Web Browsers, Office Products) *

- ◯ I install the latest version as soon as it is released
- ◯ I try to keep all software up to date when possible
- ◯ I only update software if it starts causing problems
- ◯ I don't normally update software on my machine
- ◯ I'm not sure

Do you have a firewall enabled on the device *

- ○ Yes
- ○ No
- ○ I'm not sure
- ○ What is a firewall?

Is the user account you primarily use on your device an admin user (i.e. it has permissions to install applications)

- ○ Yes
- ○ No
- ○ I'm not sure

Do you allow other users to use your device? *

- ○ Yes
- ○ No
- ○ Maybe

Do you regularly backup data on your device

- ○ Yes
- ○ No

Section 3 – Password Hygiene

How often do you change the password on your student account? *

- I have never changed the password on my student account since I got it
- I have changed my password once since I got the account
- I change my password on a regular basis
- I'm not sure

Thinking of the password you use for your student account, how long is this password? *

- 8 characters long
- Between 8 and 11 characters long
- Longer than 12 characters?
- Would rather not say

Still thinking of the password you use for your student account, how complex would this password be? (complexity can be defined as how many of the following types of characters it uses - (1) Lowercase letters, (2) Uppercase Letters, (3) Numbers, (4) Special Characters) *

- The password uses only 1 of these type of characters
- The password uses only 2 of these type of characters
- The password uses only 3 of these type of characters
- The password uses all 4 of these type of characters
- Would rather not say

How often do you generally change passwords for other accounts you use? (such as social media accounts, other email addresses etc) *

- ○ Never
- ○ Rarely
- ○ Every month
- ○ Every 3-6 months
- ○ Only if the system prompts me to
- ○ I'm not sure

Have you ever used the same password on multiple sites *

- ○ Yes, I do this all the time
- ○ Yes, but not that often
- ○ No, I use a different password for each site
- ○ Would rather not say

Would you regard the password(s) you use for your accounts as strong passwords? *

- ○ Yes
- ○ No
- ○ Maybe
- ○ I'm not sure

In relation to your online accounts (social media, email etc) Do you use a 3rd party password manager to store your passwords? *

- ○ Yes
- ○ No
- ○ I'm not sure what a password manager is

Do you allow your web browser (such as Google Chrome) to store passwords for you? *

- ○ Yes
- ○ No
- ○ I'm not sure

Do you know what Two-Factor Authentication is (also known as Multi Factor Authentication) and have you implemented this on any of your online accounts where it is offered? *

- ○ Yes, I know what it is and I have implemented it on some or all of my online accounts
- ○ Yes, I know what it is, but I have not implemented it on any of my accounts
- ○ No, I don't know what it is

Section 4 – Data Protection

Do you use a USB key (pen drive) or external hard drive to store data? *

- ○ Yes
- ○ No

Is the USB Key (pen drive) or external hard drive encrypted? *

- ○ Yes
- ○ No
- ○ I'm not sure

Section 5 – Wireless technologies

Have you ever connected to an open/insecure wireless connection (from your laptop or mobile device)?

- ○ Yes
- ○ No
- ○ I'm not sure how to tell

Have you ever checked your Online Banking or sent an email over an open/insecure wireless connection?

- ○ Yes
- ○ No
- ○ I'm not sure

Are you aware that, by using the appropriate tools, a hacker could intercept your wireless traffic if you are using an open/insecure network?

- Yes
- No
- I don't care

Q1

Quiz: In order for your home wireless network to be secure, it should have the following: (choose one)

- WPA2, unique complex password and WPS disabled
- WEP unique password, WPS enabled
- WPA2, default password, WPS disabled
- Separate 2.4Ghz and 5Ghz channels
- I have no idea

Section 6 - Quiz

Q2

You've just clicked on a web link contained in a suspicious email and now the * computer has started to behave strangely. What should you do next?

- You have security software and a firewall enabled on the device which will block malicious code from getting into yo...
- You need to update and run your anti-virus software
- Contact the IT Helpdesk or information security team
- Keep an eye on the performance of the machine

Q3

You have Multi-factor Authentication (MFA) enabled on your email account. One day, you receive an SMS which contains a one-time passcode for accessing your email account. You haven't tried to log into your email and shouldn't be getting this text. What should you do? *

- Ignore the text, it was probably a mistake
- Disable MFA which will prevent these texts from being sent to your phone
- Change your email password ASAP
- None of the above
- I honesty don't know

Q4

How often should you back up your data? *

- Once per week
- Once per month
- Once per fortnight
- Whenever you create new files or upload pictures you don't want to lose

Q5

What is the best way to validate a legitimate email vs. a phishing email *

- Bad grammar & poor spelling are the tell-tale signs of a phishing email
- Check the email headers to see where the mail came from
- Contact the sender using another medium other than email to verify if they sent the email
- A phishing email will have incorrect logos and pictures
- I have no idea

Q6

You have a highly sensitive document that you need to email to someone outside of TUDublin. What is the safest way to send this? *

- Make sure you scan the document with anti-virus software before sending it
- Send the document using HEAnet Filesender, password protect it and send the password using a different medium ...
- Send the document from an alternative email account
- Anything sent by email is encrypted by default and only the receiver can view it, so you don't have to do anything
- I have no idea

Q7

Your personal email account has been compromised. What is the best way to prevent further unauthorised access to this account? *

- Change the login password to your computer
- Update the email application software to the latest version
- Change the password and enable two-factor authentication on the email account
- Change the password and scan your computer using anti-virus software
- I have no idea

Q8

You find a USB key on the ground in the College. It is safe to plug this into your own computer to try and determine who the owner is, as long as you have your firewall enabled *

- True
- False

Q9

Your existing password for your student account is "February" and the IT department has requested you change it. Which one of the passwords below would be regarded as the most secure (i.e. harder to crack) *

- February1
- Febru@ry
- PhoneCoffeeMonitor
- September!

Q10

You have an old laptop that possibly has personal details stored on it. You want to give the laptop away to charity. What is the safest way to keep the file contents confidential? *

- Login to the machine and delete the files in question
- Replace the hard drive in the machine with a new one
- Install a new copy of the OS on the hard drive
- Login to the computer, delete the files in question AND empty the recycle bin
- I have no idea

Q11

You contact the IT helpdesk in TUDublin as you are having problems accessing your email account. The helpdesk user asks you for your password so they can test it at their end. what do you do? *

- Give them the password, as it is ok to share your password with the helpdesk
- Give them the password, but change it as soon as the phone call is finished
- Don't give them the password, call to the helpdesk in person instead

Q12

When travelling, you carry a USB stick, which is stored in your laptop bag along with your laptop, which is used to back up important information on your laptop. What is the risk here?

- ○ Electromagnetic airport scanners can corrupt the USB stick
- ○ Backups should never be stored in the same location as the original data set
- ○ Electromagnetic radiation from the laptop could corrupt the data on the USB stick
- ○ Malware from the laptop could corrupt the backup disk
- ○ I have no idea

Section 7 – Self Evaluation

Having just completed the Behaviour Analysis section of the survey, can you please now rate yourself on a scale of 1-7 what you think your level of security awareness is (1=Very Poor, 2=Poor, 3=Fair, 4=Good, 5=Very Good, 6=Excellent, 7=Exceptional)

	1	2	3	4	5	6	7	
Very Poor	○	○	○	○	○	○	○	Exceptional

Future Training

Description (optional)

Do you think Security awareness training should be provided by TU Dublin for all students?

- ◯ Yes, it should be done yearly for all students.
- ◯ Yes, it should be done for all new students
- ◯ No, I think it is up to the student to train themselves
- ◯ I dont know

APPENDIX B

This part of the document contains information relating to a chi-square "goodness of fit" test to determine if the sample obtained in the survey was representative of the actual population. It also carries out the same test on a subset of users, who own a device and opted to answer questions relating to their own password habits, to determine if this sub-set of users was representative of the survey sample.

Pearson's chi square comparison of survey results with student population

A chi-square comparison was carried out between the data gathers from the survey and compared with the actual population. To achieve this, student information was obtained from the Strategic Development Services Team within TU Dublin. This information identified the breakdown of age for each student, the college each student was located in, the status of each student in relation to being part-time or full-time.

Statistics were also obtained from the Higher Education Authority to identify the number of male and female students that had enrolled in TU Dublin for the 2017/18 Academic year. These results are outlines in the various sections below

Gender

Population		
Gender	count	percentage
Male	9233*	59.14
Female	6379*	40.86
Total	15612*	100

*These figures were obtained from the Higher Education Authority and relate to students that enrolled in the 2017/2018 academic year, excluding grauates.

Sample		
Gender	count	percentage
Male	340	45.70

Female	404	54.30
Total	744	100

We can use the Chi Square goodness of fit test to compare our sample with the actual student population using the following formula:

$$X^2 = \frac{\Sigma(O - E)^2}{E}$$

Based on the student population, 59.14% are male and 40.86% are female. Translating these percentage values to raw values based on our sample size, this gives us the figures of males = 440 and females = 304.

$$X^2 = \frac{(340 - 440)^2}{440} + \frac{(404 - 304)^2}{304}$$

By using the chi square formula, we can confirm that the Chi square value is 55.61. The p-value is <.00001. The result is significant at p <.05. We can, therefore, not accept that the sample of Male and Female respondents is representative of the student population.

Area of study

Population		
College	Count	%
COAT	4889	25.85
COB	4586	24.25
CEBE	6021	31.84
COSH	3416	18.06
Total	18912	100.00

Sample		
College	Count	%
COAT	142	18.88
COB	184	24.47
CEBE	189	25.13

COSH	237	31.52
Total	752	100

Using the same formula as above, the chi square value is 99.80, meaning the p-value is < .00001. The result is significant at $p < .05$. We can, therefore, not accept that the sample obtained in relation to area of student is representative of the student population.

Student Status (Full-time / Part-time)

Population		
Status	count	percentage
Full-Time	13835	70.85
Part-time	5693	29.15
Total	19528	100

Sample		
Status	count	percentage
Full-Time	610	81.12
Part-time	142	18.88
Total	752	100

Using a chi-square comparison between the sample population and the actual population, we get a value of 39.09. The p-value is < .00001, meaning the result is significant at $p < 0.5$. In relation to the status of the student in the sample population, we cannot accept that the sample population is representative of the full student population.

Age

A distribution of the age profile of students within TU Dublin was provided by the Strategic Development Services Team. The breakdown of the population is outlined below

Population		
Age Range		Percentage
17-19	1381	7.07

20-21	5286	27.07
22-23	4471	22.90
24-27	3276	16.78
28-34	2310	11.83
35-44	1769	9.06
45-54	787	4.03
55+	247	1.26
Total	19527	100

The breakdown of the age profile of the students that responded to the survey are outlined in the table below.

Sample		
Age Range	Count	Percentage
17-19	136	18.09
20-21	211	28.06
22-23	140	18.62
24-27	88	11.70
28-34	58	7.71
35-44	65	8.64
45-54	42	5.59
55+	12	1.60
Total	752	100

A chi-square comparison was used to determine if the sample population was representative of the actual population. The value obtained from this calculation was 163.13. The p-value is < .00001, meaning the result is significant at $p < 0.5$.

In relation to the breakdown of age values in the sample population, we cannot accept that the sample population is representative of the full student population.

Pearson's chi square comparison of subset of respondents with sample in relation to assessing device security habits and password habits

As part of the research to analyse the security habits of respondents, it was necessary to remove data where respondents had stated that they did not own a personal device or stated that they used something other than a PC or Apple Mac Laptop. Respondents that also chose "I would rather not say" when asked about their password length or if they had used the same password on multiple sites were also excluded from this analysis. After these respondents were removed from the results, a total of 590 respondents were left in the survey.

A chi-square comparison was carried out between this subset of respondents (n=591) and compared with the original number of respondents (n=752). This was carried out to determine if this subset of responses was representative of the total number of responses received.

Gender

A total of 7 respondents in this subset of data stated that they would prefer not to disclose their gender, which gave a figure of 583 total respondents. For this chi-square test, we used a significance level of 5% ($\alpha = 0.05$)

Sample		
Gender	count	percentage
Male	340	45.70
Female	404	54.30
Total	744	100

Subset		
Gender	count	percentage
Male	247	42.37

Female	336	57.63
Total	583	100

Based on the initial sample size, 45.7% are male and 54.3% are female. Translating these percentage values to raw values based on our subset size, this gives us the figures of males = 266 and females = 317.

By using the chi square formula, we can confirm that the Chi square value in this case is = 2.4959. Using a chi-square table with a DF = 1 and a significance level of 0.05, we find a critical value of 3.84, which is greater than the value found
We can, therefore, accept that the subset of male and female respondents is representative of the sample obtained from the survey.

Area of Study

Sample		
College	Count	%
COAT	142	18.88
COB	184	24.47
CEBE	189	25.13
COSH	237	31.52
Total	752	100

Subset		
College	Count	%
COAT	115	19.5
COB	148	25.1
CEBE	139	23.6
COSH	188	31.9
Total	590	100

$$X^2 = \frac{(115-111)^2}{111} + \frac{(148-144)^2}{144} + \frac{(139-148)^2}{148} + \frac{(188-186)^2}{186}$$

A chi-square comparison was used to determine if the subset was representative of the sample obtained from the survey. By using the chi square formula, we can confirm that the Chi square value in this case is = 0.823. Using a chi-square table with a DF = 3 and a significance level of 0.05, we find a critical value of 7.815, which is greater than the value found using the chi square formula.

We can, therefore, accept that the subset representing the area of study related to each respondent is representative of the sample obtained from the survey.

Level of Study

Sample		
Level of Study	Count	%
Apprenticeship / Trades	2	0.3
Undergraduate (1st year)	172	22.9
Undergraduate (year 2, 3 or 4)	432	57.4
Graduate (Masters)	125	16.6
Post Graduate (PhD)	21	2.8
Total	752	100

subset		
Level of Study	Count	%
Apprenticeship / Trades	0	0
Undergraduate (1st year)	143	24.2
Undergraduate (year 2, 3 or 4)	333	56.4
Graduate (Masters)	97	16.4
Post Graduate (PhD)	17	2.9
Total	590	100

$$X^2 = \frac{(0-2)^2}{2} + \frac{(143-135)^2}{135} + \frac{(333-339)^2}{339} + \frac{(97-98)^2}{98} + \frac{(17-17)^2}{17}$$

A chi-square comparison was used to determine if the subset was representative of the sample obtained from the survey. By using the chi square formula, we can confirm that the Chi square value in this case is = 2.59. Using a chi-square table with a DF = 4 and a significance level of 0.05, we find a critical value of 9.488, which is greater than the value found using the chi square formula.

We can, therefore, accept that the subset representing the area of study related to each respondent is representative of the sample obtained from the survey.

Student Status

Sample		
Status	count	percentage
Full-Time	610	81.12
Part-time	142	18.88
Total	752	100

subset		
Status	count	percentage
Full-Time	479	81.2
Part-time	111	18.8
Total	590	100

$$X^2 = \frac{(479-479)^2}{479} + \frac{(111-111)^2}{111}$$

A chi-square comparison was used to determine if the subset was representative of the sample obtained from the survey. By using the chi square formula, we can confirm that the Chi square value in this case is = 0. There was no difference between the sample and the subset in this case. We can therefore, accept that the subset of full-time and part-time students is representative of the sample obtained from the survey.

Age

A breakdown of the age range of each respondent is presented below.

Sample		
Age Range	Count	Percentage
17-19	136	18.09
20-21	211	28.06
22-23	140	18.62
24-27	88	11.70
28-34	58	7.71
35-44	65	8.64
45-54	42	5.59
55+	12	1.60
Total	752	100

Subset		
Age Range	Count	Percentage
17-19	116	19.7
20-21	169	28.6
22-23	102	17.3
24-27	67	11.4
28-34	42	7. 1
35-44	49	8.3
45-54	35	5. 9
55+	10	1.7
Total	590	100

$$X^2 = \frac{(116-107)^2}{107} + \frac{(169-166)^2}{166} + \frac{(102-110)^2}{110} + \frac{(67-69)^2}{69} + \frac{(42-45)^2}{45} + \frac{(49-51)^2}{51} + \frac{(35-33)^2}{33} + \frac{(10-9)^2}{9}$$

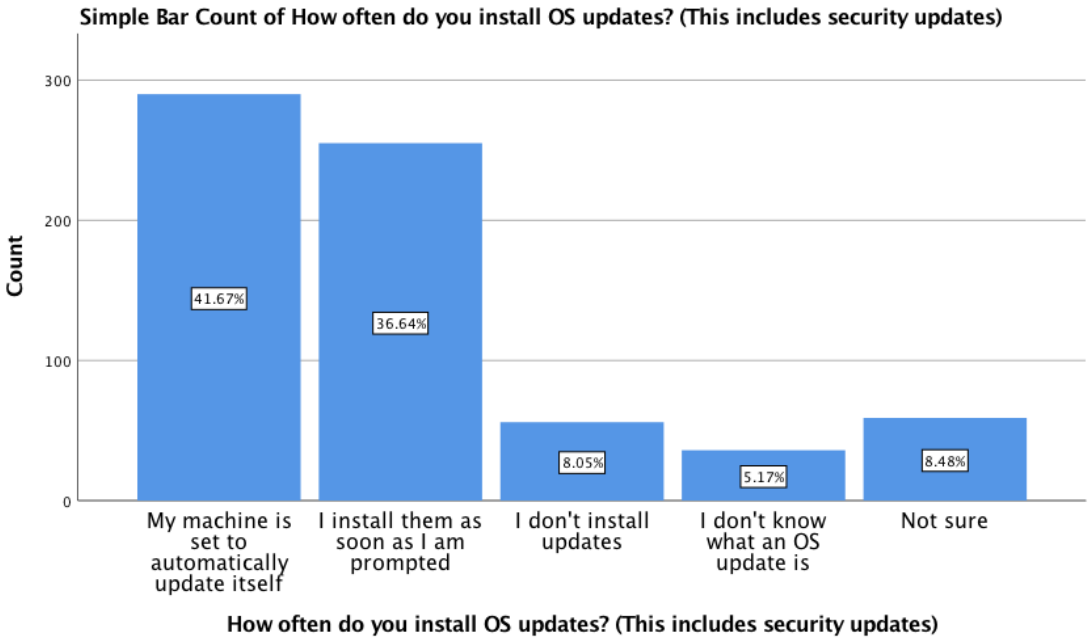
A chi-square comparison was used to determine if the subset was representative of the sample obtained from the survey. By using the chi square formula, we can confirm that the Chi square value in this case is = 1.9532. Using a chi-square table with a DF = 7 and a significance level of 0.05, we find a critical value of 14.067, which is greater than the value found using the chi square formula.

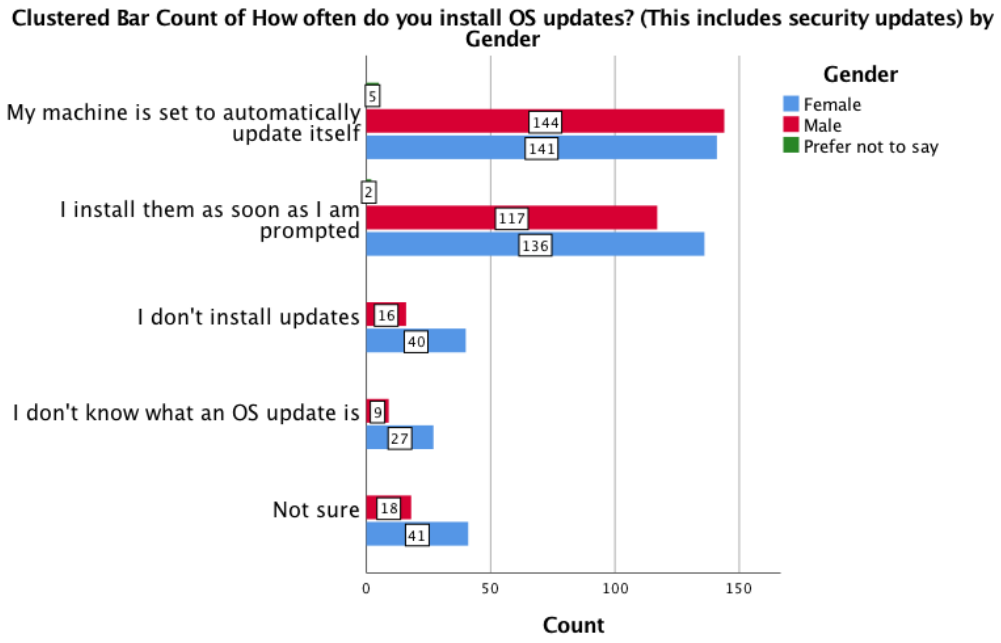
We can, therefore, accept that the subset of age ranges is representative of the sample obtained from the survey.

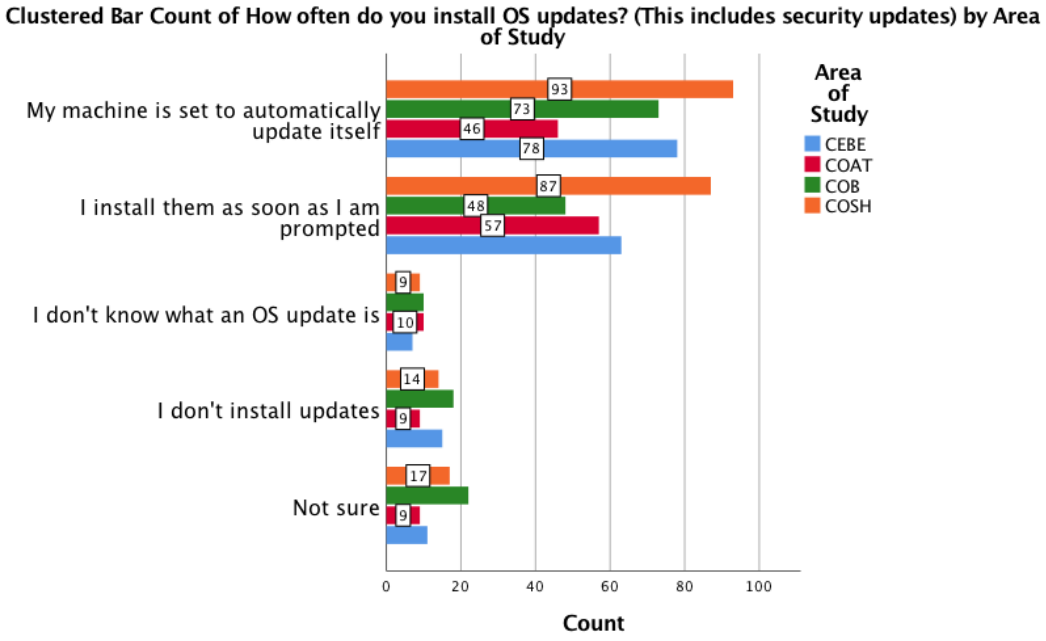
APPENDIX C

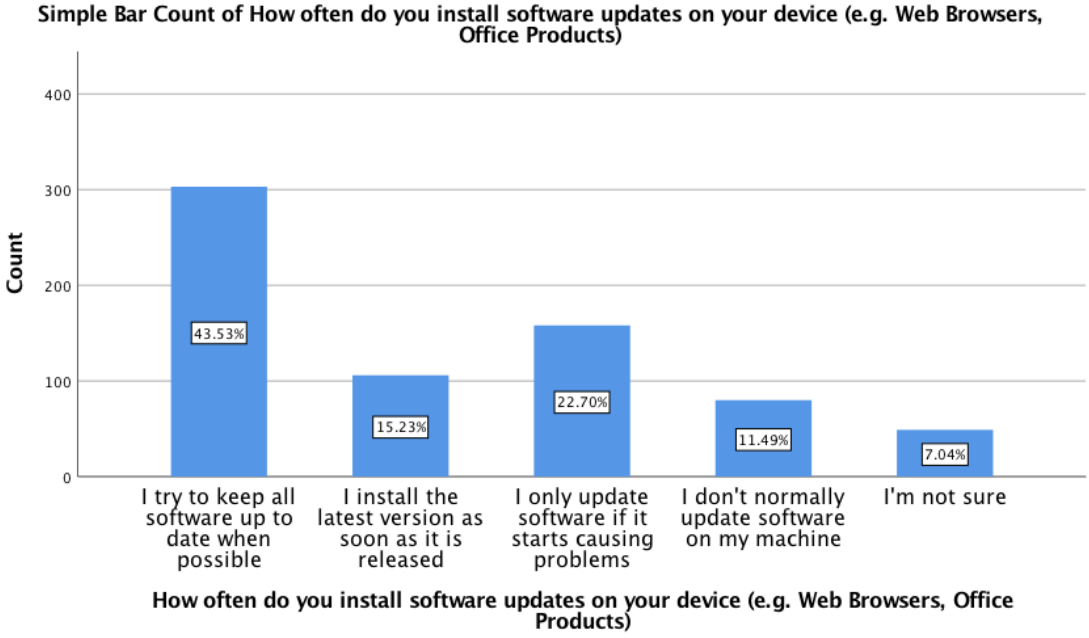
A number of additional results relating to device security, anti-virus and password habits are listed in this chapter. Not all results shown here contain any noticeable differences between the various demographic groups.

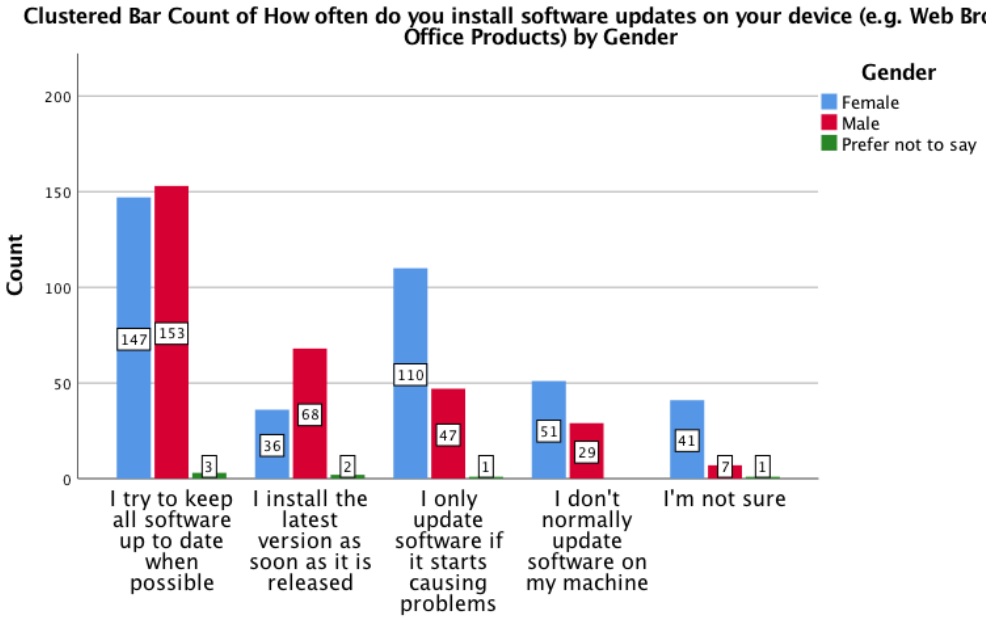
Device Security

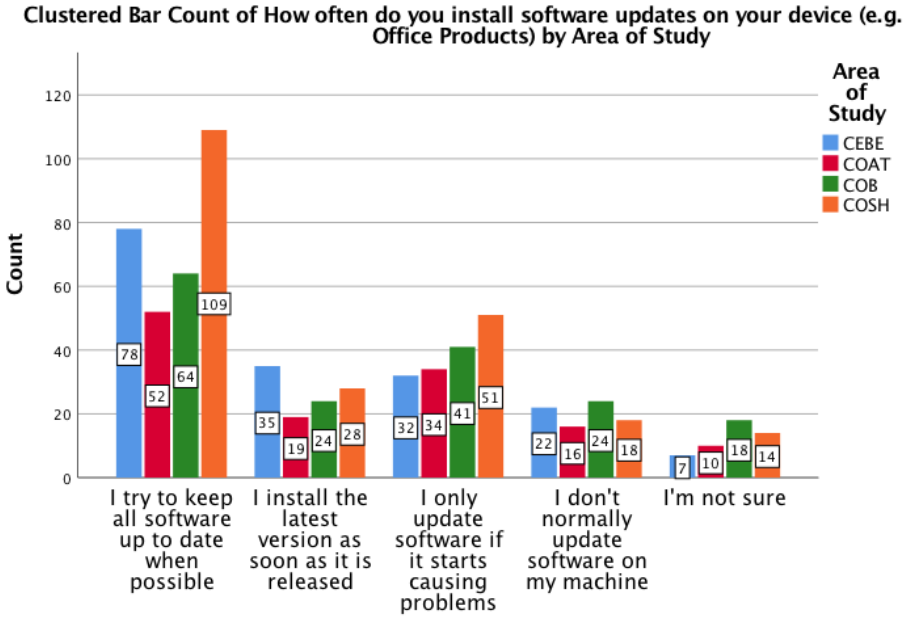


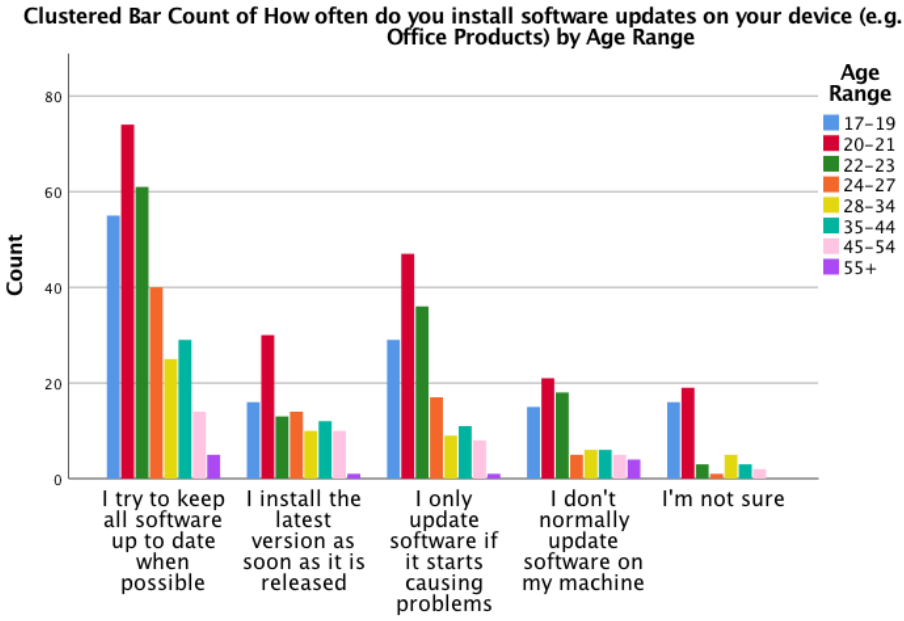


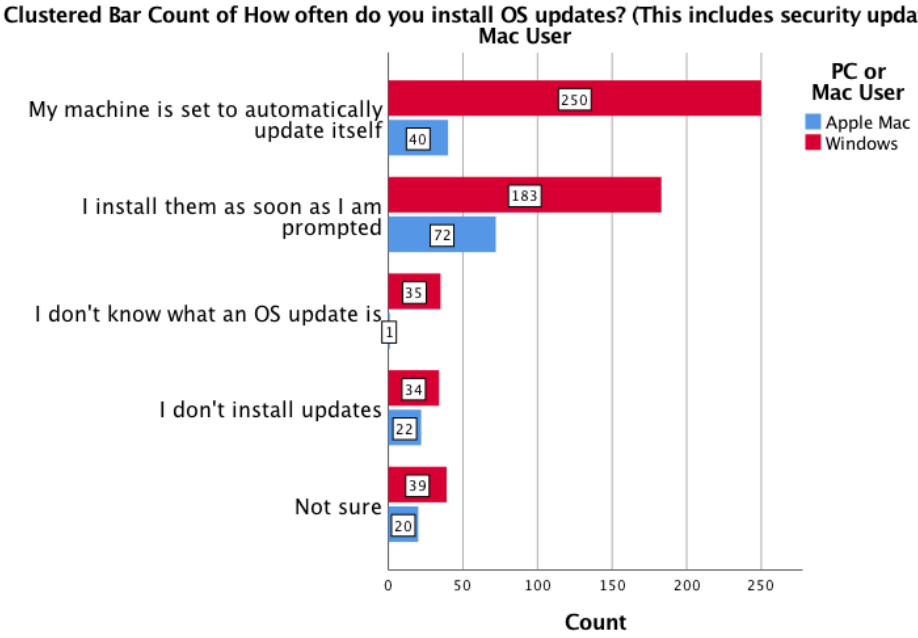

Clustered Bar Count of How often do you install OS updates? (This includes security updates) by Area of Study
My machine is set to automatically update itself
I install them as soon as I am prompted
I don't know what an OS update is
I don't install updates
Not sure
93
73
46
78
87
48
57
9
10
14
9
17
9
0
20
40
60
80
100
Count
Area of Study
CEBE
COAT
COB
COSH


Simple Bar Count of How often do you install software updates on your device (e.g. Web Browsers, Office Products)
Count
400
300
200
100
0
43.53%
15.23%
22.70%
11.49%
7.04%
I try to keep all software up to date when possible
I install the latest version as soon as it is released
I only update software if it starts causing problems
I don't normally update software on my machine
I'm not sure
How often do you install software updates on your device (e.g. Web Browsers, Office Products)

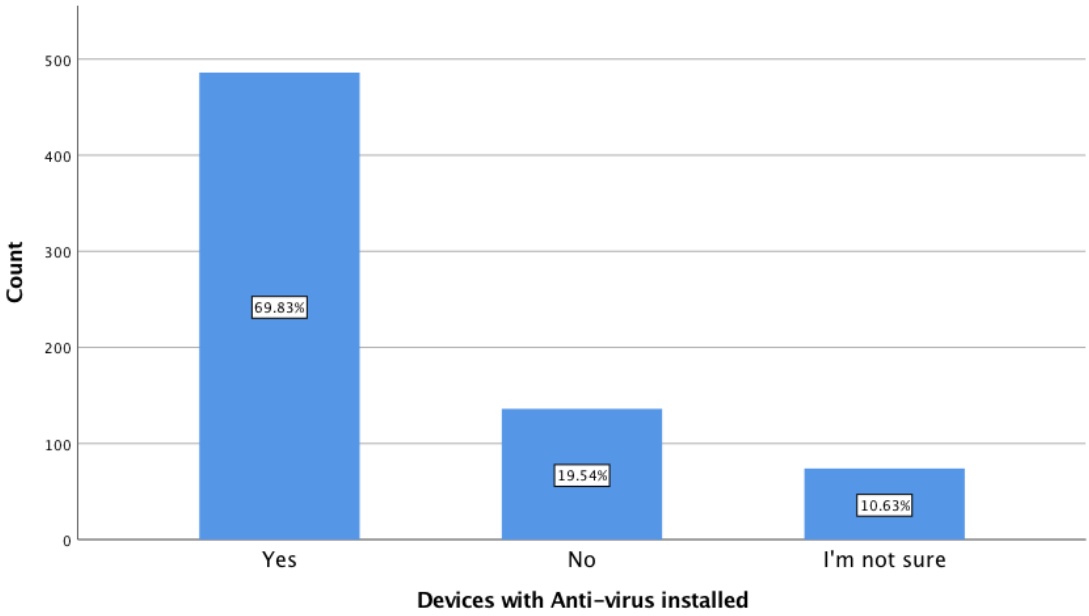

Clustered Bar Count of How often do you install software updates on your device (e.g. Web Browsers, Office Products) by Gender
Gender
Female
Male
Prefer not to say
Count
200
150
100
50
0
147
153
3
36
68
2
110
47
1
51
29
41
7
1
I try to keep all software up to date when possible
I install the latest version as soon as it is released
I only update software if it starts causing problems
I don't normally update software on my machine
I'm not sure

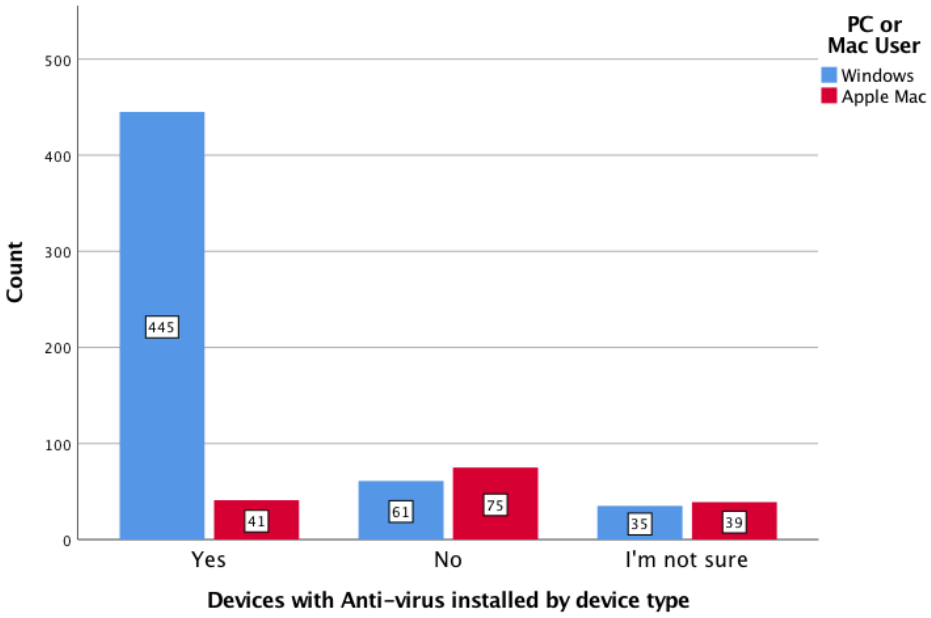

Clustered Bar Count of How often do you install software updates on your device (e.g. Web Browsers, Office Products) by Area of Study
Area of Study
CEBE
COAT
COB
COSH
Count
120
100
80
60
40
20
0
78
52
64
109
35
19
24
28
32
34
41
51
22
16
24
18
7
10
18
14
I try to keep all software up to date when possible
I install the latest version as soon as it is released
I only update software if it starts causing problems
I don't normally update software on my machine
I'm not sure


Clustered Bar Count of How often do you install software updates on your device (e.g. Web Browsers, Office Products) by Age Range
Count
80
60
40
20
0
Age Range
17–19
20–21
22–23
24–27
28–34
35–44
45–54
55+
I try to keep all software up to date when possible
I install the latest version as soon as it is released
I only update software if it starts causing problems
I don't normally update software on my machine
I'm not sure

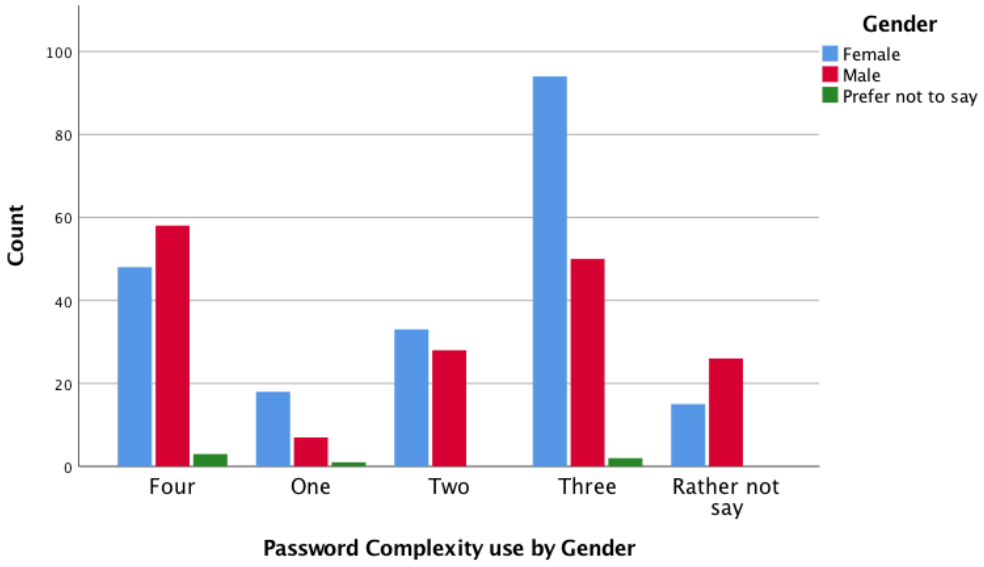

Clustered Bar Count of How often do you install OS updates? (This includes security updates) by PC or Mac User
My machine is set to automatically update itself
I install them as soon as I am prompted
I don't know what an OS update is
I don't install updates
Not sure
250
40
183
72
35
1
34
22
39
20
0
50
100
150
200
250
Count
PC or Mac User
Apple Mac
Windows

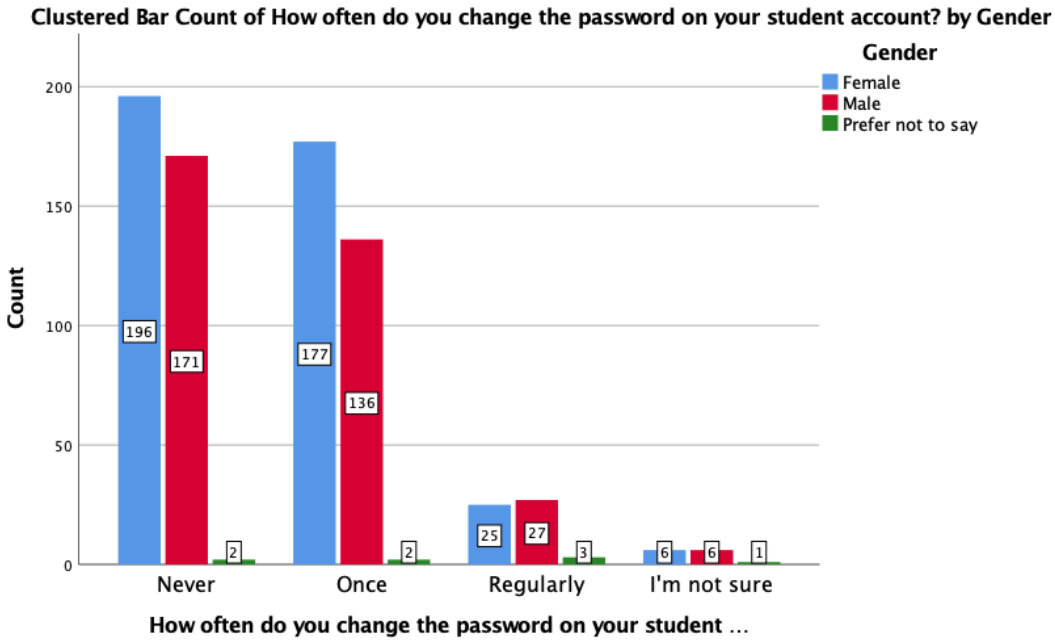
Anti-virus / Anti-Malware installed

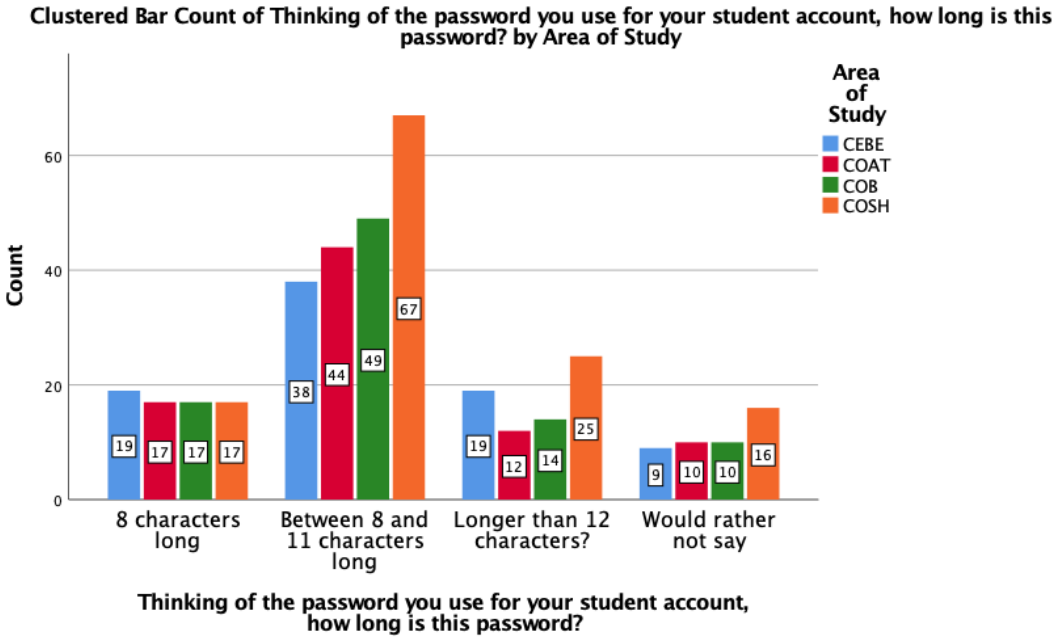


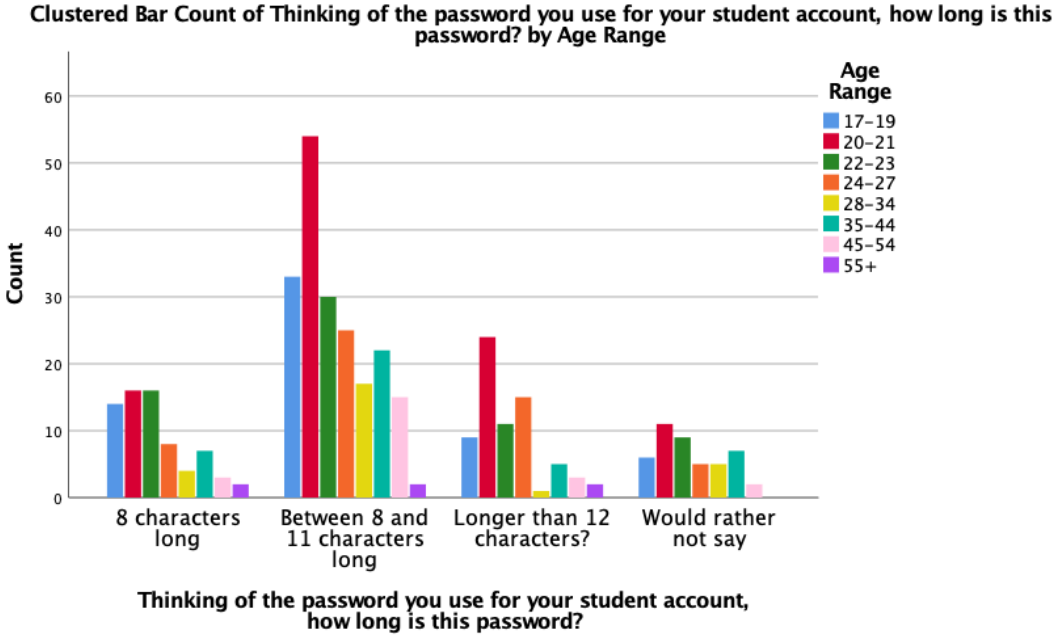


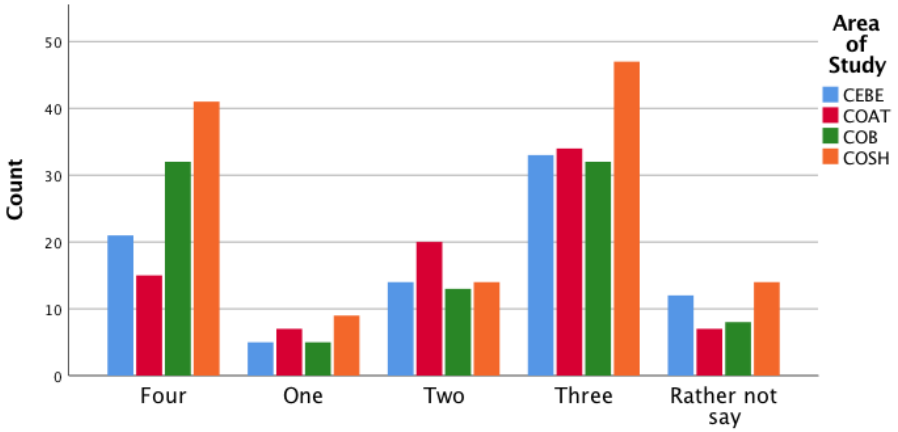
Password Statistics



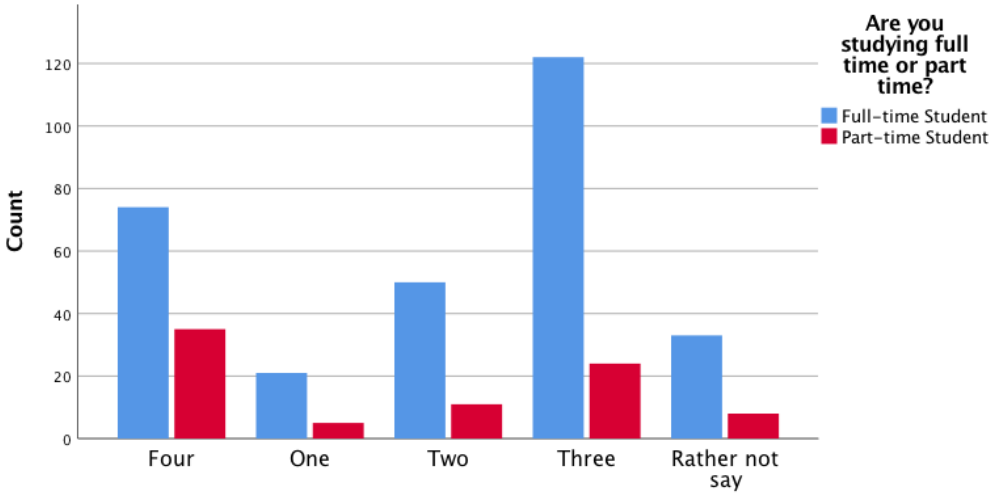



Clustered Bar Count of Thinking of the password you use for your student account, how long is this password? by Area of Study
Area of Study
CEBE
COAT
COB
COSH
Count
0
20
40
60
19
17
17
17
38
44
49
67
19
12
14
25
9
10
10
16
8 characters long
Between 8 and 11 characters long
Longer than 12 characters?
Would rather not say
Thinking of the password you use for your student account, how long is this password?

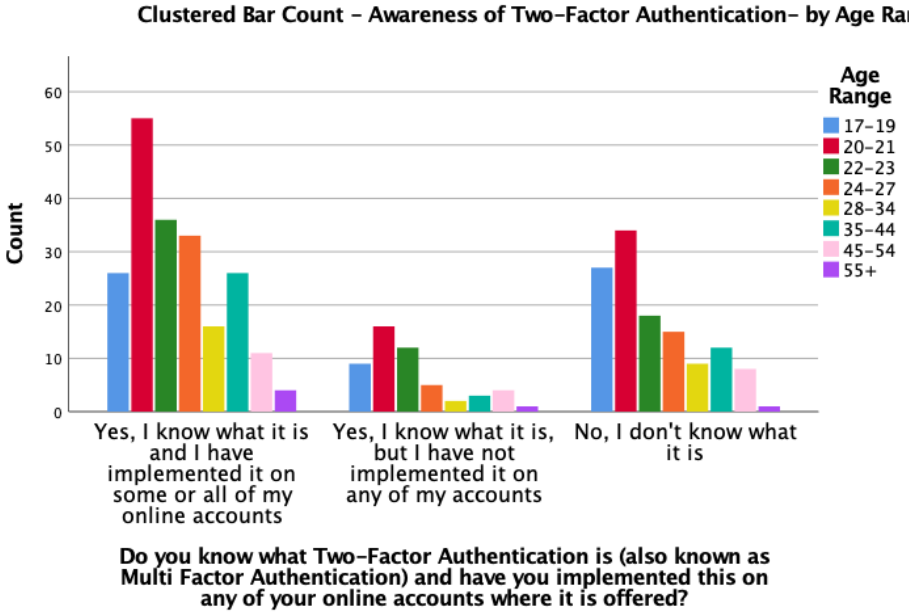

Clustered Bar Count of Thinking of the password you use for your student account, how long is this password? by Age Range
Age Range
17–19
20–21
22–23
24–27
28–34
35–44
45–54
55+
Count
0
10
20
30
40
50
60
8 characters long
Between 8 and 11 characters long
Longer than 12 characters?
Would rather not say
Thinking of the password you use for your student account, how long is this password?

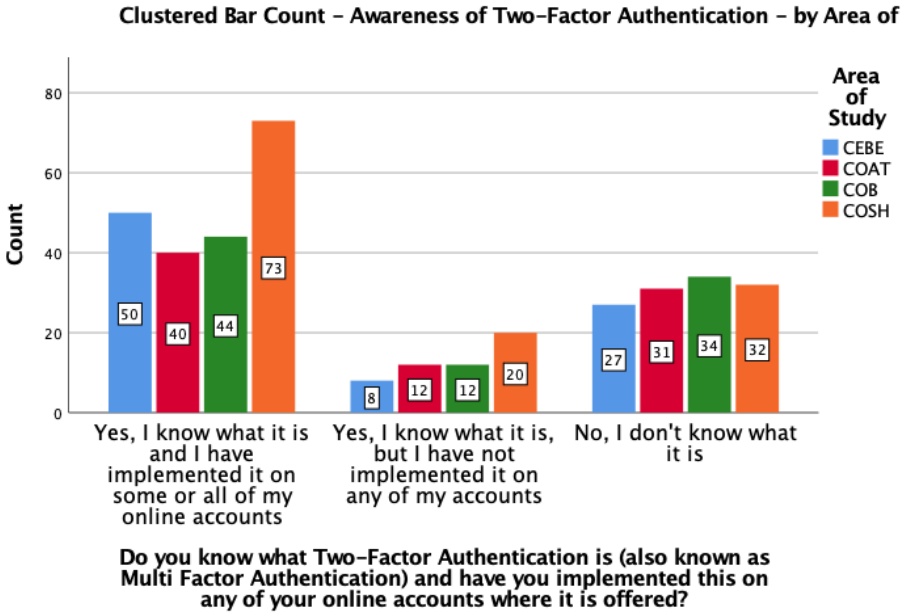


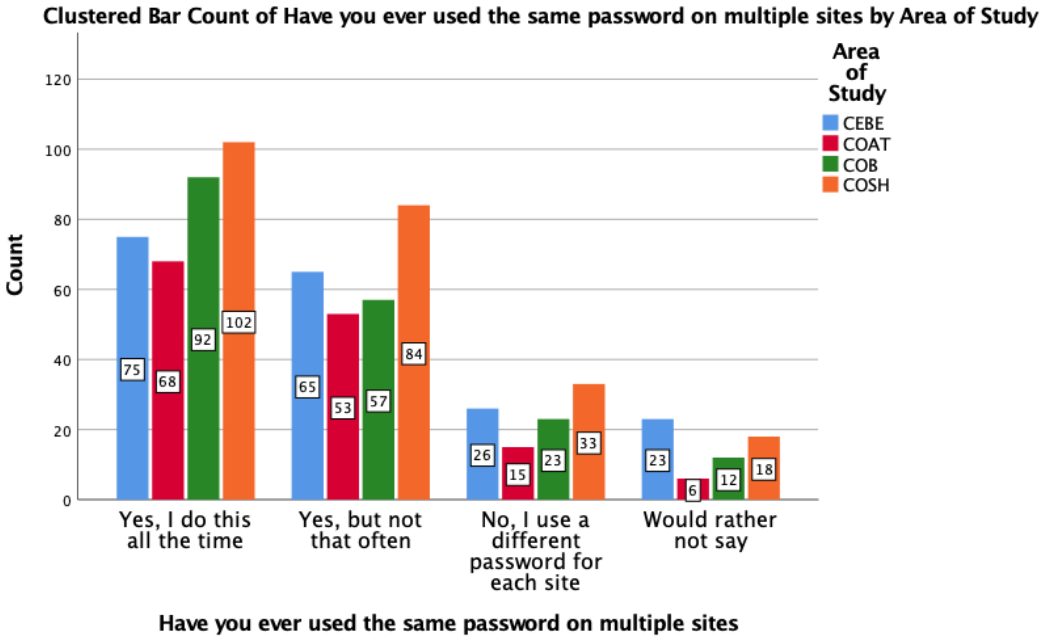
Password Complexity – Breakdown by Area of Study (CEBE = College of Engineering and Built Environment, COAT = College of Applied Arts and Tourism, COB = College of Business, COSH = College of Science and Health)



Password Complexity use by Student Status


Clustered Bar Count - Awareness of Two-Factor Authentication- by Age Range
Count
60
50
40
30
20
10
0
Age Range
17-19
20-21
22-23
24-27
28-34
35-44
45-54
55+
Yes, I know what it is and I have implemented it on some or all of my online accounts
Yes, I know what it is, but I have not implemented it on any of my accounts
No, I don't know what it is
Do you know what Two-Factor Authentication is (also known as Multi Factor Authentication) and have you implemented this on any of your online accounts where it is offered?


Clustered Bar Count - Awareness of Two-Factor Authentication - by Area of Study
Count
80
60
40
20
0
Area of Study
CEBE
COAT
COB
COSH
50
40
44
73
8
12
12
20
27
31
34
32
Yes, I know what it is and I have implemented it on some or all of my online accounts
Yes, I know what it is, but I have not implemented it on any of my accounts
No, I don't know what it is
Do you know what Two-Factor Authentication is (also known as Multi Factor Authentication) and have you implemented this on any of your online accounts where it is offered?


Clustered Bar Count of Have you ever used the same password on multiple sites by Area of Study
Area of Study
CEBE
COAT
COB
COSH
Count
120
100
80
60
40
20
0
75
68
92
102
65
53
57
84
26
15
23
33
23
6
12
18
Yes, I do this all the time
Yes, but not that often
No, I use a different password for each site
Would rather not say
Have you ever used the same password on multiple sites

APPENDIX D

The following breakdown of student population was provided by the Strategic Development Services Team within TU Dublin – City Centre campus on the 16th of April 2019 by Mark Russell, which was based on Data available from March 2019. These figures exclude incoming exchange students but includes apprentices. This gives exact figures of Age, Student Status and the College the student is enrolled in. Gender breakdown was not provided by the University.

Age

AGE	TOTAL
17	3
18	124
19	1254
20	2563
21	2724
22	2595
23	1876
24	1200
25	854
26	649
27	573
28	487
29	403
30	362
31	293
32	271
33	248
34	246
35	206
36	245
37	202
38	202

AGE	TOTAL
46	98
47	105
48	86
49	86
50	73
51	70
52	62
53	52
54	44
55	47
56	31
57	32
58	25
59	26
60	22
61	13
62	12
63	7
64	10
65	8
66	4
67	2

39	185
40	181
41	137
42	144
43	143
44	124
45	111

68	1
69	2
70	1
71	2
79	1
81	1
Grand Total	19528

Status

MODE	TOTAL
FT	13835
PT	5693
Grand Total	19528

College

COLLEGE	TOTAL
Arts & Tourism	4889
Business	4586
Engineering & Built Environment	6021
Graduate Research School	539
LTTC	77
Sciences & Health	3416
Grand Total	19528