A Demographic Analysis to Determine User Vulnerability among Several Categories of Phishing Attacks.

Robert Griffin
Technological University Dublin

Follow this and additional works at: https://arrow.tudublin.ie/scschcomdis

Part of the Computer Sciences Commons

Recommended Citation

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License
A demographic analysis to determine user vulnerability among several categories of phishing attacks.

Robert Griffin
C11473422

A dissertation submitted in partial fulfilment of the requirements of Dublin Institute of Technology for the degree of M.Sc. in Computing (Security and Forensics)

June 2018
Declaration

I, ROBERT GRIFFIN, certify that this dissertation which I now submit for examination for the award of M.Sc. 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 text of my work.

This dissertation was prepared according to the regulations for postgraduate study of the Dublin Institute of Technology 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 Institutes guidelines for ethics in research.

Signed: ____________________________

Date: 15th June 2018
Abstract

Phishing attacks have been on a meteoric rise in the last number of years, with 2016 seeing a 65% increase. The attacks range from targeting individuals with personalised messages to spam attacks from bot accounts. With the chances of being targeted by a phishing attack increasing, it is important to identify who is most at risk in order to help alleviate this threat.

The aim of this study is to examine members from several demographics and their vulnerability to three types of phishing using data collected from a survey (n = 198). The survey tested the participant’s ability to recognise spoofed phishing emails, SMS phishing (Smishing) and content spoofing attacks. The respondents were presented with questions in the form of screenshots using real world phishing examples. Their answers were collected which recorded whether they got each question correct or incorrect. The data collected was analysed using a two sample t-test or one-way Anova depending on the number of categories per demographic.

This study addressed demographic vulnerability to different types of phishing and highlighted who is most at risk. The results of the research revealed that gender and income did not play a part in a participant’s vulnerability to phishing when analysing their total scores across each type of phishing. However, age, education and occupation presented statistically significant results to indicate they do.

Keywords: Phishing; Demographics; Vulnerability; Survey
Acknowledgement

I would firstly like to thank my project supervisor, Patrick Tobin, for his guidance and support throughout this research project. His input and advice was greatly appreciated at every stage and it helped guide my progress.

To my parents, Gerry and Catherine, and the rest of my family for their encouragement, support and patience.

Finally, a massive thank you to my girlfriend Laura who was always there for me with advice, kind words and a willingness to proof read at any time. To my friends who helped me along the way to distribute the survey and to Alex who gave me invaluable statistical advice.
# Table of Contents

Abstract ........................................................................................................................................ ii
Acknowledgement ....................................................................................................................... iii
1 Introduction ............................................................................................................................... 1
   1.1 Background ......................................................................................................................... 1
   1.2 Research Problem ............................................................................................................... 2
      1.2.1 Research Questions ..................................................................................................... 3
   1.3 Research Objectives ......................................................................................................... 4
   1.4 Research Methodologies ................................................................................................... 5
   1.5 Scope and Limitations ....................................................................................................... 5
   1.6 Document Outline ............................................................................................................ 6
2 Literature Review and Related Work ...................................................................................... 7
   2.1 Introduction ....................................................................................................................... 7
   2.2 What is Phishing? .............................................................................................................. 7
   2.3 Why We Fall for Phishing Attacks ................................................................................... 8
   2.4 Anatomy of an Attack ...................................................................................................... 9
   2.5 Types of Phishing Attacks ............................................................................................... 10
      2.5.1 Email ......................................................................................................................... 10
      2.5.2 Spear Phishing .......................................................................................................... 11
      2.5.3 Content Spoofing ...................................................................................................... 12
      2.5.4 Smishing ................................................................................................................ 12
      2.5.5 Vishing .................................................................................................................... 13
   2.6 Demographics .................................................................................................................. 13
      2.6.1 Age and Gender ....................................................................................................... 14
      2.6.2 Education ............................................................................................................... 14
      2.6.3 Income .................................................................................................................. 15
      2.6.4 Occupation/Student ............................................................................................... 15
2.7 Dealing with Phishing ................................................................. 16
  2.7.1 Approaches to Examining the Problem .................................. 16
  2.7.2 Methods to Counteract Phishing ........................................... 16
2.8 Conclusion .................................................................................. 17
3 Design and Methodology ................................................................. 18
  3.1 Introduction .............................................................................. 18
  3.2 Design Overview ..................................................................... 18
  3.2.1 Demographic Breakdown .................................................... 19
  3.3 Survey Design and Responses .................................................. 20
  3.3.1 Why a Survey was used ....................................................... 20
  3.3.2 Design .............................................................................. 22
  3.3.3 Respondent Population ........................................................ 23
  3.3.4 Preamble ........................................................................... 24
  3.4 Statistical Tools and Methods Used .......................................... 25
  3.4.1 Variables ........................................................................... 25
  3.4.2 Two-sample T Test ............................................................... 25
  3.4.3 One-way Anova ................................................................. 26
  3.4.4 Statistical Software ............................................................. 28
  3.5 Conclusion .............................................................................. 28
4 Implementation and Results ............................................................. 30
  4.1 Introduction .............................................................................. 30
  4.2 Survey Responses ..................................................................... 30
  4.2.1 Completion Rates ............................................................... 30
  4.2.2 Demographic Distribution ................................................... 30
  4.3 Descriptive Statistics ................................................................. 34
  4.3.1 Gender .............................................................................. 34
  Gender – Total Scores ................................................................. 34
6.5 Future Work & Recommendations .................................................................57
6.6 Final Thoughts.......................................................................................58
Appendix A - Survey Questions ....................................................................59
   Section 1 – Demographic Information.......................................................59
   Section 2 – Phishing Questions.................................................................61
Appendix B - Descriptive Statistics...............................................................71
Gender........................................................................................................71
   Gender – Total Scores............................................................................71
   Gender – Category 1 Scores ....................................................................71
   Gender – Category 2 Scores ....................................................................71
Age.............................................................................................................72
Education....................................................................................................74
   Education – Total scores........................................................................74
Income........................................................................................................75
Occupation/Student....................................................................................77
7 Bibliography .............................................................................................79
Table of Figures

Figure 1-1: Responses to the current threat of phishing .......................................................... 2
Figure 2-1: Number of fraudulent attempts to steal data .......................................................... 11
Figure 3-1: Histogram to identify the normality of data ............................................................ 28
Figure 4-1: Percentile distribution of gender ............................................................................. 31
Figure 4-2: Percentile distribution of age groups ....................................................................... 31
Figure 4-3: Percentile distribution of education levels ............................................................... 32
Figure 4-4: Percentile distribution of annual gross income brackets ........................................ 33
Figure 4-5: Percentile distribution of sectoral occupation/students .......................................... 33
Figure 4-6: The percentage of correct answers per question in the survey ............................... 38
Figure 4-7: T-test formula ........................................................................................................ 40
Table of Tables

Table 4-1: Descriptive statistics of the total test scores for gender ........................................35
Table 4-2: Descriptive statistics of the total test scores for age ..................................................35
Table 4-3: Descriptive statistics of the category 3 test scores for age ..................................35
Table 4-4: Descriptive statistics of the total test scores for education ..................................36
Table 4-5: Descriptive statistics for the category 1 test scores for education ..................36
Table 4-6: Descriptive statistics of the total test scores for the income demographic ..........36
Table 4-7: Descriptive statistics of the category 1 test scores for the income demographic ...37
Table 4-8: Descriptive statistics of the total test scores for the occupation/student demographic ....................................................................................................................................37
Table 4-9: Descriptive statistics of the category 1 test scores for the occupation/student demographic ....................................................................................................................................37
Table 4-10: Descriptive statistics of the category 3 test scores for the occupation/student demographic ....................................................................................................................................38
Table 4-11: A breakdown of the demographic categories with coding applied ..................39
Table 4-12: Two sample t-test values, which are to be analysed ........................................40
Table 4-13: Values for the one-way Anova and post hoc test which are to be analysed ......40
Table 4-14: Analysis of results among gender regarding the four test scores ....................42
Table 4-15: Statistically significant results for the age demographic when analysing their total test scores .....................................................................................................................................43
Table 4-16: Analysis of results among age regarding the four test scores .......................45
Table 4-17: Statistically significant results for the education demographic when analysing their total test scores .....................................................................................................................................46
Table 4-18: Analysis of results among education regarding the four test scores ...............47
Table 4-19: Category 1, 2 and 3 scores when running a one-way Anova on gross income brackets ....................................................................................................................................48
Table 4-20: Analysis of results among income regarding the four test scores ...............48
Table 4-21: Statistically significant results for the occupation demographic when analysing their category 3 test scores .....................................................................................................................................50
Table 4-22: Analysis of results among occupation/students regarding the four test scores ....51

Table B-1: Descriptive statistics of the total test scores for gender ..........................................71
Table B-2: Descriptive statistics of the category 1 test scores for gender .............................71
Table B-3: Descriptive statistics of the category 2 test scores for gender ............................ 71
Table B-4: Descriptive statistics of the category 3 test scores for genders .......................... 72
Table B-5: Descriptive statistics of the total test scores for age ........................................ 72
Table B-6: Descriptive statistics of the category 1 test scores for age ................................. 72
Table B-7: Descriptive statistics of the category 2 test scores for age ................................ 73
Table B-8: Descriptive statistics of the category 3 test scores for age ................................ 73
Table B-9: Descriptive statistics of the total test scores for education ............................... 74
Table B-10: Descriptive statistics of the category 1 test scores for education ..................... 74
Table B-11: Descriptive statistics of the category 2 test scores for education ...................... 74
Table B-12: Descriptive statistics of the category 3 test scores for education ...................... 75
Table B-13: Descriptive statistics of the total test scores for income ................................. 75
Table B-14: Descriptive statistics of the category 1 test scores for income ......................... 76
Table B-15: Descriptive statistics of the category 2 test scores for income ......................... 76
Table B-16: Descriptive statistics of the category 3 test scores for income ......................... 77
Table B-17: Descriptive statistics of the total test scores for occupation ............................. 77
Table B-18: Descriptive statistics of the category 1 test scores for occupation .................... 77
Table B-19: Descriptive statistics of the category 2 test scores for occupation .................... 78
Table B-20: Descriptive statistics of the category 3 test scores for occupation .................... 78
1 Introduction

1.1 Background

As technology advances, there is a greater reliance on the internet for people to conduct business and socialise. This advancement in technology has also resulted in cybercriminals following the trend as there is an increased risk of users being the target of fraud (Downs, Holbrook, & Cranor, 2007). One of the main types of internet fraud is phishing which “is a form of social engineering attack in which phishers, i.e. attackers, trick the victim to fraudulently obtain private information” (Darwish, Zarka, & Aloul, 2013). The information collected by these attacks “can be used in identity theft, to remove funds from a customer account, and in theft of online resources” (Wardman, 2016). Phishing has been a growing problem for internet users and organisations as methods to deceive users and gain sensitive information have become more commonplace and sophisticated (Anti-Phishing Working Group, 2017). Furthermore, the Anti-Phishing Working Group’s 2016 report observed that “2016 ended as the worst year for phishing in history with 1.2 million attacks, 65% more than 2015”. As phishing becomes more of a problem, the reasons why people fall victim to these exploits are examined. The psychology behind why the attacks succeed and who fall for them need to be further explored.

As part of the initial research for this project, a survey study of twenty-three cyber security professionals was conducted. The survey examined whether anti-phishing training and education could prevent attacks against organisations. The results, as shown in Fig.1-1, found that 78% of the respondents believed phishing “is one of the most significant threats we face today.” (Griffin, 2017).

---

1 This study was conducted in 2017 using Survey Monkey. The population was sourced through the author’s employer. It was sent to the cyber security and risk advisory department.
1.2 Research Problem

This study will highlight the similarities between certain demographics and their vulnerability to three phishing categories. Researchers have studied why people fall for phishing attacks however; limited research has been done to study demographic factors in vulnerability to phishing. By determining which groups are most vulnerable to phishing, we can potentially outline the best course of action to prevent and protect against it. (Sheng, Holbrook, Kumaraguru, Cranor, & Downs, 2010).

The objective of this research project is to identify differences that exist among several demographics to determine which population group is more vulnerable to being exploited by a particular phishing category. There are many different types of phishing such as email, search engine, session hijacking, DNS-based and deceptive phishing. Although email is widely used, “phishing has spread beyond email to include VoIP, SMS, instant messaging, social networking sites, and even massively multiplayer games” (Hong, 2012).

Three categories of phishing have been selected for this research project due to their recurring nature and potential to cause harm (Hong, 2012; Huang, Zhong, & Tan, 2009). These include:

1. Email – Phishes sent via email
2. Smishing – SMS phishing where an attacker uses text messages to exploit victims.
3. Content Spoofing – Altering some of the content of a reliable page to trick users.

Five demographic groups were chosen in order to compare and contrast their vulnerability to the three phishing categories. The demographics chosen for this study are:

1. Gender
2. Age
3. Education
4. Income
5. Occupation/Student

1.2.1 Research Questions
To determine whether demographics have an influence on a person’s vulnerability to phishing, the research seeks to answer:

RQ1: Are members of demographic groups more vulnerable to being exploited by a particular category of phishing attack?

With the research question defined, several hypotheses were outlined which are to be investigated:

Hypothesis 1:

H_{0}: When given a quiz, there is no difference in the mean test scores between males and females as to who are more vulnerable to fall for a phishing attack.

H_{1}: When given a quiz, there is a difference in the mean test scores between males and females as to who are more vulnerable to fall for a phishing attack.

Hypothesis 2:

H_{0}: When given a quiz, there is no difference in the mean test scores between age groups as to who are more vulnerable to fall for a phishing attack.

H_{1}: When given a quiz, there is a difference in the mean test scores between age groups as to who are more vulnerable to fall for a phishing attack.
Hypothesis 3:

**H₀**: When given a quiz, there is no difference in the mean test scores between education levels as to who are more vulnerable to fall for a phishing attack

**H₃**: When given a quiz, there is a difference in the mean test scores between education levels as to who are more vulnerable to fall for a phishing attack

Hypothesis 4:

**H₀**: When given a quiz, there is no difference in the mean test scores between income levels as to who are more vulnerable to fall for a phishing attack

**H₄**: When given a quiz, there is a difference in the mean test scores between income levels as to who are more vulnerable to fall for a phishing attack

Hypothesis 5:

**H₀**: When given a quiz, there is no difference in the mean test scores between occupations or students as to who are more vulnerable to fall for a phishing attack

**H₅**: When given a quiz, there is a difference in the mean test scores between occupations or students as to who are more vulnerable to fall for a phishing attack.

1.3 Research Objectives

Most of the past work in the area of phishing has focused on why people fall for an attack, how to deal with detection, trends, and the economic impact. Little research has been carried out regarding what type of users are vulnerable to being phished (Kumaraguru et al., 2009a; Sheng et al., 2010). After performing a thorough review of the relevant literature, it appears no study has been undertaken to determine what types of phishing are the most successful at exploiting these users.

This research study is a quantitative evaluation of vulnerability among different types of phishing. Whether gender, age, education, income or occupation influences the user’s vulnerability to phishing is determined by this study. The objective of the research is to administer a scored survey to capture demographic information of a large and diverse population group. The survey will include questions relating to three phishing categories that users undertake. The results will be analysed to determine if the different users defined by their demographics are more vulnerable than others are to being phished by one or more of the categories.
1.4 Research Methodologies

The research methodologies used in this study included primary research collecting demographic information and participants answers from a scored survey on the three defined categories of phishing attacks. A survey format was the chosen research method for this study, utilised to acquire responses from the population group. The survey was administered online so the target population could be reached more easily and the results could be obtained quicker. The format of the survey presented each question with mandatory status, providing respondents with pre-defined answers. This ensured that each participant completed the survey fully, and prevented the risk of partial responses which would render the answers unusable.

The participants were presented with five questions relating to each category of phishing. Each question asked the respondent if it is legitimate or not i.e. would they use and interact with the content as if they normally would. Quantitative research was carried out by statistically analysing if a difference exists between the mean scores of each category among the demographic groups. The quantitative methodology was selected because of the types of data analysis that are needed in order to accurately draw conclusions from the data gathered. Empirical analysis of the data was carried out to either accept or reject the defined hypotheses. An inductive approach was used to develop theories on the vulnerability of demographic groups per phishing category.

1.5 Scope and Limitations

The literary research for this study covers many topics, from phishing to how and why people fall victim to phishing attacks and historically what demographics are most vulnerable. The assessment of vulnerability to a particular type of phishing was limited to three that were viewed as the most prevalent and dangerous. There are wide ranges of different phishing attacks, and it would be unachievable to assess each type given the scope of this research.

The survey will attempt to assess demographic vulnerability in respect to the three types of phishing identified. It would be preferable to assess a large number of demographics but given the wide range of different demographics that exist, and the scope of this research project, five were selected. The limitations of this study include such areas as non-response bias. Those who choose not to respond to the survey could have different views and phishing awareness than those who do. Given the number of categories per
demographic group, which will divide the amount of responses across the group, a large sample size is needed if the results are to obtain a high statistical power. If low response numbers are recorded for some of the categories within a demographic group, it will be hard to infer robust results. One aim of this research was to project the results onto a larger population. With generalisability in mind, it may be hard to do this without a proper random sample obtained.

1.6 Document Outline

- **Chapter 2**: This consists of the literature review, which will outline what phishing is, how attacks are performed and the different types of phishing. Demographic vulnerability to phishing and how to counteract it are also presented.

- **Chapter 3**: This chapter outlines the design and methodology of the study. It focuses on why a survey was used for data collection and the statistical tools and methods used to analyse the data.

- **Chapter 4**: This chapter illustrates the distribution of the demographic groups as well as the significant descriptive statistics that were recorded. Finally, the statistical analysis is carried out that will conclude whether or not the null hypotheses are rejected.

- **Chapter 5**: This discusses the results obtained from Chapter 4 in light of what is already known. The results are used to confirm or refute previous work and highlight what new findings have emerged.

- **Chapter 6**: This gives a short account of the results obtained and the research problem that was being addressed. Limitations of the study are highlighted and any future work that might be undertaken is recommended.
2 Literature Review and Related Work

2.1 Introduction

In this chapter, a number of topics are discussed that will present an overview of phishing and some of the important points relating to how attacks are carried out and demographic vulnerability. Firstly, a description of the origins of phishing and why people fall for phishing attacks are introduced. The anatomy of how phishing attacks are usually performed and the different types of phishing that exist are discussed. Furthermore, analysis of the literature that discusses demographic vulnerability to phishing and approaches on how to deal with and counteract phishing are presented.

This chapter’s purpose is to set out and identify the concepts that relate to the research question and to highlight any gaps in the literature that may exist between phishing and demographic vulnerability.

2.2 What is Phishing?

Phishing is a type of social engineering attack in which cybercriminals use fake or spoofed emails and deceitful web sites in order to trick people into giving out sensitive information. Social engineering is “any act that influences someone to take an action that may or may not be in his or her best interest” (Hadnagy, 2014, p. 27). The victims of these attacks perceive the emails and websites as legitimate but they are the work of criminals looking to perform fraud and identity theft (Sheng & Magnien, 2007). Although the exact origins of phishing may be open to debate, Symantec report that the first instances of phishing attacks occurred in the mid 1990’s targeting America Online (AOL) (Symantec, 2007). They go on to say “The attackers typically used either instant messages or email to trick users into divulging their AOL passwords. Victims would provide the attackers with this information, which the attackers would, in-turn, leverage to assume ownership of the victim’s AOL account”.

Parno, Kuo, & Perrig (2006), concluded that phishing was a growing problem that threatens businesses and consumers due to its exploitative nature. It is widely seen as the most common form of fraud to steal personal information (Eisenstein, 2007; Keith B. Anderson, Erik Durbin & Salinger, 2008). The damage caused is not just financial as trust from customers is also affected. Due to the phish usually appearing as if it is from a reputable source, the reputation of the business might also be damaged in the process. Consumers may associate the negative effects of the fraud with the company and might
stop using their services completely. According to a Gartner survey in 2007, “3.6 million U.S. adults lost a total of 3.2 billion dollars directly due to phishing” (Huang et al., 2009). Phishing attacks usually begin by targeting the public and aiming to steal personal information but it has also evolved to target more high profile victims. This is known as whaling, a process where senior executives within a business are targeted in the hopes they reveal sensitive company information or corporate secrets.

2.3 Why We Fall for Phishing Attacks

Many users fall victim to phishing attacks because they lack the knowledge to properly identify them (Sheng et al., 2010). Cybercriminals prey on people’s vulnerabilities and take advantage of this by attempting to phish the victim using fear tactics, strict deadlines or stressful situations. Research shows that there are several reasons why people are vulnerable to phishing attacks (Sheng et al., 2010). The authors state “people tend to judge a website’s legitimacy by its look and feel which attackers can easily replicate”.

There are a number of tools at the disposal of cybercriminals to clone a reputable web site such as LinkedIn or Facebook. The Social Engineer Toolkit and httrack using Kali Linux can clone webpages and potentially trick the user into entering personal information (Anoop, 2016). Some of the earliest studies investigating why people fall for phishing attacks were undertaken by (Dhamija, Tygar, & Hearst, 2006; Downs, Holbrook, & Cranor, 2006; Downs et al., 2007; Wu, Miller, & Garfinkel, 2006).

- Dhamija et al., (2006), states, “22 participants were shown 20 web sites and asked to determine which ones were fraudulent”. They found that 90% of the participants were fooled by the best phishing sites, with few noticing the security indicators presented to them.

- Wu, Miller, & Garfinkel (2006) commented, “Users fail to continuously check the browser’s security indicators, since maintaining security is not the user’s primary goal”. The authors attribute this to the users relying on the web content to decide if the site is authentic or not. This supports the point from Sheng et al., (2010), clearly showing that users do not spend much time judging whether the site is real or spoofed outside of the content displayed to them.
• Third, the perceived consequences are not a good predictor of identifying or avoiding a phishing attack. Downs, Holbrook, & Cranor (2007) state, “The ratings of consequences suggest that fear of credit card theft is not a great motivator for protecting one’s information”. The authors go on to note, “protections against phishing might not gain much traction from warnings about how easy it would be for a phisher to steal one’s card.”

• Lastly, although users may be aware of phishing, it does not reduce their vulnerability to identifying an attack as they employ their own methods to decipher whether the content is real or spoofed (Downs et al., 2006). Downs et al., (2006), go on to say “Participants used various strategies to make decisions about the trustworthiness of email, mostly centred around interpreting the text of the email rather than any more reliable cues in headers or URLs associated with links”.

2.4 Anatomy of an Attack
A phishing attack may seem like an easy, juvenile fraud but there are layers to it that require expertise and knowledge to exploit people’s natural tendencies. According to Hong (2012), there are three major phases to a phishing attack.

1. The first is for potential victims to receive a phish. Most phishing scams are sent via email as they can be quickly produced and mass distributed. As previously mentioned, the phish usually uses social techniques, rather than technical tricks to fool end users. The author remarks, “Conveying urgency is a well-known method used by criminals to misdirect people’s attention. An example is pretending to be a system administrator warning people about a new attack, urging them to install the attached patch.”

2. The second is the victim performing the suggested action in the message. Once the victim has incorrectly flagged the phishing attempt as real, they are usually asked to do something such as install software or follow a hyperlink. Most phishing attacks try to convince people to go to a spoofed website where personal and sensitive information will be collected. “The user connects to a spoof website by clicking on a link in the email. Their web browser may access the website directly or be redirected from an initial site to the actual phishing pages” (Moore
The scammers will usually try to host a fake site using a compromised machine or register a new domain. When registering a new domain, the scammers look for names similar to the site they want to impersonate, and employ homograph attacks\(^2\). These are done to “exploit the visual similarity of characters, for example, bankofthevvest.com uses two v’s to look like a w” (Hong, 2012).

3. Finally, the third is the criminal monetising stolen information. Once the user has been presented with an accurate clone of the target site, they are usually requested to fill in their personal information. “The compromised details are usually emailed to a webmail address, but are sometimes stored in plain text files at the spoof website, awaiting direct collection by the fraudster” (Moore & Clayton, 2007). Once the details are received, they are usually then sold on to cashiers who will target areas such as bank accounts. Moore and Clayton (2007), go on to say, “The mean lifetime of a normal phishing site is 61.69 hours”, illustrating how quickly this type of fraud can happen and disappear without a trace.

2.5 Types of Phishing Attacks

This study will look to demonstration the potential vulnerabilities certain demographics have in relation to three categories of phishing attack. Email is the most common form of phishing, capable of many attack vectors such as deceptive phishing, malware attacks and DNS based attacks (Jones, Towse, & Race, 2015).

2.5.1 Email

Due to the reliance on email by most of the people in the world for either business or personal use, the number of potential targets for cybercriminals is massive. Sending an e-mail and asking for a user’s bank account login details is a simple idea and it costs almost nothing. “Each day more and more e-mails are sent with the aim of making the web users believe that the same is legitimate and from the trusted institutions” (Purkait, 2012). Email phishing can take place in many ways, each with their own level of creativity and deception. One of the most common methods employed to trick a user is to convey a sense of urgency, such as notifying the victim they have failed to login

\(^2\) A homograph is a letter or string that is visually confusable with a different letter or string - https://www.gribble.org/papers/usenix06_homograph.pdf
multiple times and must verify their account details or risk severe consequences (Hong, 2012).

As aforementioned in Chapter 1.1, research was conducted using a survey prior to this study with twenty-three cyber security professionals, examining whether anti-phishing training and education could prevent attacks against organisations. Of the twenty-three participants, twelve stated that they have received a phishing email in the last six months intending to steal their personal information as illustrated in the figure below. A further ten respondents remarked that they have received both an email and phone call from an unreliable source trying to steal their personal information (Griffin, 2017).

![Bar chart showing the number of fraudulent attempts to steal data](image)

**Figure 2-1: Number of fraudulent attempts to steal data**

### 2.5.2 Spear Phishing

Spear phishing is a more sophisticated form of email phishing, which is utilised to target individuals, small groups or organisations. Zhao, An, & Kiekintveld (2016), state that it “uses personal information and social engineering to craft very believable messages with the goal of inducing the recipient to open an attachment, or visit an unsafe website by clicking a link.” Although similar to regular email phishing, spear phishing is a much more costly process than sending spam email messages in the hopes of an everyday user interacting with it. For example, if a cybercriminal is looking for specific data that an organisation holds, they must first find out who has access to the data and target those people specifically (Wardman, 2016).
2.5.3 Content Spoofing

Content spoofing refers to an attack when a criminal inserts malicious content into a legitimate site in order to trick the user into thinking it is real. This malicious content can “redirect to other sites, install malware on a user’s computer, or insert a frame of content that will redirect data to a phishing server” (Jakobsson, Myers, 2006, p36). These types of attacks are made possible by an injection vulnerability in the web application that fails to correctly handle user-supplied data. It is possible for cybercriminals to lure a victim into viewing spoofed content by inserting malicious links into emails, web pages or forums. Such spoofing attempts can be hindered if the users are trained and know what to look for within the content of the page. Kong (2011), states, “However, a hacker may modify the information and the links in an established website by altering the content on the server of a legitimate organisation. This mode of content spoofing is more difficult to detect because the casual Internet user cannot tell whether the page is legitimate or not.”

2.5.4 Smishing

A mobile device is defined as a small computing device, usually with a small output screen, which may have touch input or miniaturised keyboard (W. C. Hu, Y. Zuo, 2009). In 2016, it was estimated that over 62% of the world’s population owned a mobile phone, with that number expecting to exceed five billion by 2019\(^3\). Today, the use of mobile phones is almost a necessity in order to browse the internet or use the devices applications; be it for business or personal use (Parasuraman, 2017). Due to phishing being on a dramatic rise, it comes as no surprise that phishing via landline and mobile phones is now a big concern. Instead of a phishing attack happening through email or a spoofed webpage, smishing relies on short messaging services (SMS) or text messages (Yeboah-Boateng & Amanor 2014). Yeboah-Boateng & Amanor (2014), expresses there are two main processes for smishing scams. The first “involves receiving a text message which is purported to have originated from a known and trusted source, such as your bankers or your system administrator”. The second “involves you receiving a vital text message about your identity being stolen or account number being frozen”. As

---

previously mentioned in chapter 2.3, the same human instincts of fear, stress and severe consequences are preyed upon to have the user fall for this fraud. The SMS or text message then directs you to a website via hyperlink or a phone number so you can verify your details. Upon receiving this sensitive information, the cybercriminals use it to withdraw money from the victim’s bank account or attempt to set up a new credit card with their name. Smishing also has the added danger of attachments being sent with the original message, which the victim thinks is from a reputable source. This can lead to “the attachment downloading a virus or malware unto the victim’s device which in turn installs a root kit or backdoor for the scammers to have access to everything on the phone” (Yeboah-Boateng & Amanor, 2014).

2.5.5 Vishing

Vishing is another form of phishing involving phone calls, where the attacker attempts to lure the victim into providing personal information and exploit this information for personal gain. The term is derived from two different words, voice and phishing. Attackers tend to favour Voice over Internet Protocol (VoIP) rather than traditional telephone phone calls, as the whole process can be brought up and down in a short time. Noblis (2007, p. 42), gives a description of a typical vishing attack, where “an attacker calls a large number of telephone numbers, spoofing their Caller ID credentials to appear to be originating from a credit card company or some other trusted provider. They direct consumers to a telephone number that they control to collect personal information to be used for fraud, often using a voice response system.”

2.6 Demographics

“Previous work suggests that users’ demographics are useful indicators in identifying the most vulnerable users to phishing attacks” (Mohebzada, Zarka, Bhojani, & Darwish, 2012). There have been some studies completed to analyse how vulnerable certain demographics are to phishing attacks (Gavett et al., 2017a; Kumaraguru et al., 2009b; Sheng et al., 2010). The studies that have been carried out to analyse demographic vulnerability to phishing have revolved around the user’s age or gender. There appears to be gaps in the knowledge when the research involves other demographics such as income, education and occupation. This became evident during the literature collection process due to the lack of peer-reviewed research available.
2.6.1 Age and Gender
Research has been carried out to examine why phishing occurs but only a few studies look at the demographic factors relating to phishing vulnerability. Kumaraguru et al., (2009) undertook a real world phishing study to investigate demographic factors to phishing susceptibility. The project which comprised of 515 participants concluded that there was no significant difference among gender in the tendency to fall for phishing emails. The 18-25 age group was consistently seen as the most vulnerable age group to phishing attacks. Similarly, Sheng et al., undertook a study in 2010 to examine the relationship between demographics and phishing susceptibility. The authors performed a roleplay phishing exercise to analyse these factors and view the effectiveness of anti-phishing educational material. They found that “women are more susceptible than men to phishing” and supporting Kumaraguru et al., “participants between the ages of 18 and 25 are more susceptible to phishing than other age groups” (Sheng et al., 2010). These findings suggest that younger individuals are more vulnerable to a phishing attack. This could be due to a number of reasons such as a lack of education in the subject matter or being too trust worthy of web content presented to them. In a study by Gavett et al., (2017), the authors sought to determine whether older adults are more vulnerable to phishing attacks. They commented that “compared to younger adults, older adults were more likely to have knowledge of phishing and more likely to have been victimised by phishing in the past”. This is an interesting point as it could explain why younger individuals are more vulnerable to phishing. Gavett et al., (2017), further backs up this claim by remarking that in relation to phishing attacks, “being an older adult was associated with 3.69 time’s greater odds of being suspicious relative to younger adults”.

2.6.2 Education
In a study by Dhamija et al., (2006), 22 participants were shown 20 web sites and asked to determine which ones were fraudulent. They concluded that, “Good phishing websites fooled 90% of participants”. The authors found that there was “no significant correlation between education level and scores.” They also stated, “Among our participants, we did not observe a relationship between scores and sex, age, educational level or experience. A larger study is needed to establish or rule out the existence of such effects in the general population.”
2.6.3 Income

Little research appears to have been undertaken to analyse whether income is a factor in vulnerability to phishing. It makes sense that cybercriminals would target victims with more assets in order to maximise their rewards but the literature suggest differently. Leukfeldt (2014) investigates phishing victims, and the increased or decreased risk of victimisation. Leukfeldt studies whether victims have significantly higher income and financial assets than non-victims and the results show that financial characteristics do not seem to play a role. He goes on to say, “An unemployed person on a shoestring budget or a director of a multinational company: everyone has an equal chance of becoming a victim. There seems to be no evidence for so-called spear-phishing attacks on specific targets with lots of money.”

2.6.4 Occupation/Student

In contrast to income and education, there has been some research performed to look at vulnerability to phishing regarding of whether the user is a student or employed. Unfortunately, most of this research is not peer reviewed and reported on by organisational reports. In a study by Mohebzada et al., (2012), the authors performed two real life large scale phishing attacks on 10,917 members of a university including students and faculty. The experiments involved “sending spoofed emails which seemed to come from legitimate sources” which were intended to trick the subjects into revealing their confidential information. The author’s found that 9% of the enrolled students and alumni fell for the phishing attack. Comparing the results between students and faculty, it emerged that “students were more prone to phishing attacks…which indicates that experience could be a factor in victimising an individual during a social engineering attack.” Similarly, in a study by Metzger, Flanagin, & Zwarun (2003), the authors recorded both students and non-students ability to evaluate potentially spoofed online content. The results showed that “students generally found all sources of information to be more credible than the non-students did, rating all channels except the Internet as significantly more credible as did the non-students.” These findings illustrate that students are more trustworthy of online content presented to them and could potentially be more vulnerable to a phishing attack.

McAfee reported in their August, 2014 threat analysis that “employees working at accounting, finance and HR performed the worst at recognising phishing attempts, while
these are departments that arguably hold some of the most sensitive data of the organisation” (McAfee, 2014).

2.7 Dealing with Phishing

2.7.1 Approaches to Examining the Problem

Due to the rapid growth of phishing and its potential impact, several methods are usually undertaken to examine user vulnerability. These include either a questionnaire or survey to collect user data and some sort of visual exam to test the participant’s knowledge and reactions (Anti-Phishing Working Group, 2017; Butler, 2007; Downs et al., 2006; Erkkilä, 2011; Jansson & Von Solms, 2013; Kumaraguru et al., 2009b; Sheng et al., 2010). Such approaches are typically undertaken due to their proficiency for collecting and analysing user data as the tests can be done in a safe ‘lab’ environment. Other real world examples are employed such as sending out phishing emails to an entire student campus or organisation and measuring the success rate (Rocha Flores, Holm, Nohlberg, & Ekstedt, 2015). These are often the best methods to capture real phishing statistics but include their own ethical and privacy issues.

2.7.2 Methods to Counteract Phishing

The most successful techniques to counteract phishing attacks rely on interactive demos, online training materials, testing and situated learning in order to outline where these attacks occur and how to identify them (Jensen, Dinger, Wright, & Thatcher, 2017; Lastdrager, Gallardo, Junger, & Hartel, 2017; Mayhorn & Nyeste, 2012). Jensen, Dinger, Wright, & Thatcher (2017), state that “to prevent phishing attacks, organisations often rely on three techniques”:

1. “Automated removal or quarantine of phishing messages and corresponding websites”.

2. “Automated warning mechanisms that notify individuals when they encounter a suspicious message or website”.

3. “Behavioural training during which individuals are taught to identify and report attacks”.

Training materials are usually published by security firms, non-profit organisations and government entities in an effort to inform the public and organisations of the potential
danger. Griffin (2017) states that “in a paper by Sheng & Magnien (2007), the authors design and evaluate the implementation of an anti-phishing online game, Anti-Phishing Phil.” Learning science principles were employed to design and iteratively refine the game, which was evaluated through a user study. They found that “the participants who played the game were better able to identify fraudulent web sites compared to the participants in other conditions.” Training and education appear to be among the best techniques in order to mitigate against phishing attacks. Similarly to Sheng & Magnien (2007), another study was undertaken by Kumaraguru et al., (2009b), to test the long-term retention of anti-phishing training by users. Five hundred and fifteen participants were split evenly among three groups, which were control, single training and multiple training. The participants were all sent a series of legitimate and illegitimate spear phishing emails over the course of one month and the two training groups received a series of anti-phishing training material in this time. The results of the study show that:

1. “Users trained with PhishGuru retain knowledge even after 28 days.”
2. “Adding a second training message to reinforce the original training decreases the likelihood of people giving information to phishing websites.”

2.8 Conclusion

In this chapter, a variety of literature relating to phishing was examined. Firstly, an outline of phishing, its origins and how it has become a serious threat to businesses and consumers was considered. A discussion on why users fall for phishing attacks was introduced and four key points were highlighted as the primary contributing factors. The anatomy of phishing attacks were identified, discussing the three main phases that usually occur when an attack is carried out. Furthermore, several of the most common types of phishing were explored, detailing how they are used to exploit vulnerable users and the constant adaptation that they utilise in order to remain dangerous and undetected. Additionally, demographic vulnerability to phishing was reviewed from what is already known. Age, gender, income, education and occupation were selected as the five demographics to be analysed and the apparent gaps that exist in the literature relating to them were identified. Finally, the approaches to dealing with phishing and the most effective methods to counteract phishing attacks were examined.
3 Design and Methodology

3.1 Introduction
This chapter will outline the design and methodology used in order to carry out the study. It begins by giving a brief overview of the design and why new information was needed in order to complete the study. A breakdown of each demographic is described and the reasons for choosing the categories that make up the demographic groups. The reasons for choosing a survey as an adequate data collection mechanism are explained and why a convenience sample was used rather than a random sample. A more detailed explanation of the design is outlined; breaking down the survey into its two sections and mentions how the respondent population was targeted. Finally, the statistical tools and methods used to analyse the data are presented and any necessary expectations of the data are explained.

3.2 Design Overview
As aforementioned in Chapter 1, little research has been completed to study demographic factors relating to the vulnerability of phishing. A thorough search of the relevant literature was conducted and it appears that none have been undertaken linking these demographic factors with vulnerability to specific categories of phishing. Due to this, new information needed to be gathered to allow an up to date study to be completed. Although there are many different types of phishing, three have been selected due to their recurring nature and potential to cause harm (Hong, 2012; Huang et al., 2009). The three types of phishing that will be analysed in this study are:

1. Email
2. SMiShing
3. Content Spoofing

The first stage of this study was to prepare a survey that investigates how vulnerable users are to the three defined categories of phishing attacks. The first section of the survey collected demographic information so the participants can be gathered into their respective demographic groups. The second section contained five questions relating to each of the three phishing categories described above.
3.2.1 Demographic Breakdown

In order to accurately track demographic information of the respondents, each demographic set had several categories. For gender, the respondents could be male or female and to study age, the demographic set was divided into the following groups:

1. 18-25
2. 26-35
3. 36-45
4. 46-55
5. 56+

Educational categories were chosen using the European Qualifications Framework (EQF) due to this study not being exclusively limited to Ireland. It is used as “an overarching qualifications framework, which would serve as a translation device to make qualifications more readable and understandable across different countries and systems in Europe” (EFQ, 2009). The EFQ utilises eight levels, which are categorised by a set of descriptors to indicate the relevant learning qualifications. The first four levels were discarded, as the study is not concerned with the participant’s earlier achievements such as the Junior or Leaving Certificate. The lowest education category a respondent could select was Level 5, which is the equivalent of an associate degree. This degree is awarded by colleges upon completion of a course that usually lasts two years. Level 6 consists of a Bachelor’s degree or a higher diploma while level 7 is a Master’s degree or postgraduate diploma. Level 8 refers to a Doctorate degree or higher doctorate and is the highest level of education that can be chosen by participants. Finally, an option for “None of the above” was added if the participant did not have any of the outlined degrees. The options participants could select from are highlighted below:

1. Level 5 (Advanced/Higher Certificate)
2. Level 6 (Bachelor’s Degree/Higher Diploma)
3. Level 7 (Master's Degree/Postgraduate Diploma)
4. Level 8 (Doctorate Degree/Higher Doctorate)
5. None of the above

Income was also tracked between several brackets to reflect the respondent’s current yearly earnings. This information coupled with the other responses could be too intrusive for some participants, so the option to not disclose their income was added.
1. € 0 - 20,000
2. € 20,000 - 30,000
3. € 30,000 - 40,000
4. € 40,000 - 50,000
5. € 50,000 - 60,000
6. € 60,000 +
7. I don't want to disclose

Finally, the respondents would then answer whether they were a student or working, and if so, what industry they work in. Due to the large number of industries that people can work in, several of the most all-encompassing were chosen in order to get accurate results, which were:

1. Agriculture
2. Business/Finance
3. Education
4. Healthcare
5. Information Technology
6. Manufacturing
7. Media
8. Student

3.3 Survey Design and Responses

3.3.1 Why a Survey was used

The term ‘survey’ refers to the selection of a relatively large sample of people from a pre-determined population. This is usually followed by a collection of data from the individuals with which the researcher makes an inference about the wider population known as the population of interest (Kelley, Clark, Brown, & Sitzia, 2003). This study was undertaken with the hopes that the results could be projected onto a larger population set, with generalisability in mind. Some of the advantages of using a survey given the nature of this study are noted by (Kelley et al., 2003).

1. “The research produces data based on real world observations (empirical data)”.
   This is true as the phishing samples used in the survey come from real world phishing attacks. This allowed the participants to see first-hand how deceptive and legitimate phishing attacks have become.
2. “Surveys can produce a large amount of data in a short time for a fairly low cost”. Due to the project life cycle of this study only lasting several months, it was concluded that a survey was an ideal research strategy in order to capture representative results quickly.

3. “The breadth of coverage of many people or events means that it is more likely than some other approaches to obtain data based on a representative sample, and can therefore be generalisable to a population”. This is yet another reason why a survey was selected because it is hoped that this research will accurately project the results to a larger population set.

Piloting of the survey was necessary so the respondents can easily understand the concept, the questions presented and the instructions displayed. Piloting is a process of testing the research strategy/tool on a sample of members from the target population and is helpful in highlighting whether sufficient responses are available to the participants. It is important that when conducting a pilot, that the same procedure be used as if it was the real survey to outline any potential problems with responses.

The study set out to be able to project the results onto a wider population utilising a large random sample to gather responses. Brant, Haas-Haseman, Wei, Wickham, & Ponto (2015), remark that using this method will “increases the likelihood that the responses from the sample will accurately reflect the entire population.” The authors go on to say, “in order to accurately draw conclusions about the population, the sample must include individuals with characteristics similar to the population.” For this study, it was improbable to accurately obtain a random sample given the time constraints and cost it could incur. For example, gathering population information from third party sources such as a company could incur a cost. It may also take a significant amount of time to retrieve this information with no guarantee of completion. Due to this, a convenience sample was used which is a type of nonprobability sample that obtains responses in any way possible. Etikan (2016), states it is “where members of the target population that meet certain practical criteria, such as easy accessibility, geographical proximity, availability at a given time, or the willingness to participate are included for the purpose of the study.” It may be difficult to make valid inferences about the larger population group because this is probably not a true random sample. In order to mitigate against
this, a snowball sample was also employed which asks the respondents to suggest another person or group of people willing to participate in the study. With the amount of demographic information being collected from the respondents and the large number of categories per demographic set, it is hoped that it will be feasible to accurately project the results onto a larger population.

Another way to mitigate against the potential lack of generalisability is to increase the sample size. There is no definitive answer as to how many responses are necessary for a survey and this study is an analysis of quantitative data. Due to this, responses needed to be high in order to give a better estimate of the population.

3.3.2 Design

The online survey was created using Google Forms for several reasons such as the ability to easily distribute it among multiple social media platforms and its built-in statistics and analytics. In order to accurately test the participants against the types of phishing attacks that are present today, several data sources were utilised to get the samples used in the survey. Personal email accounts, mobile phones and online phishing databases such as Phishtank4 were used to collect the samples. Many different samples were identified as potential options to be used in the survey. The samples, which were chosen to be included in the survey, were selected due to their ability to look like a genuine web page or request. This was done to give a lasting impression of how dangerous and concealed phishing can be. The survey was split into two sections, which are as follows:

Section 1

Section 1 collected demographic information from five key areas, which were:

1. Gender
2. Age
3. Education
4. Income
5. Occupation/Student

As described above in chapter 2, it is clear that some demographic information has been analysed in previous studies to show how vulnerable users are to phishing. From the research, age and gender are the main demographic factors which have been previously

---

4 https://www.phishtank.com/developer_info.php
studied. In order to obtain new data on what other demographics are most vulnerable, five were selected to extend the research.

Section 2

Section 2 then presented the participants with questions. The questions were divided between the three phishing categories as follow:

- Question 1 – 5: Email Phishing
- Question 6 – 10: Smishing
- Question 11 – 15: Content Spoofing

The questions were in the form of screenshots using real world phishing examples taken from the sources listed above. These images were harmless depictions of real phishing examples and legitimate sites, which was important from an ethical point of view. Each category contained several real phishing examples and several genuine requests, which tested how accurately the participants could identify the phishing attacks. The participants answered whether the information presented to them was legitimate or not, assessing if they would interact with the content like normal. A binary scoring system of zero and one was used per question, zero indicating an incorrect answer and one indicating a correct answer. This left the participants with three separate scores out of five, relating to each of the three phishing categories and a total score out of fifteen.

3.3.3 Respondent Population

The research studied members of multiple demographic groups who participated in the survey. Surveying multiple demographic groups with no limiting restrictions allowed for greater participation due to the larger number of possible respondents. In order to gather responses to the survey, it was distributed electronically to third party social and professional groups. The survey was not distributed via internal corporate channels; this removed the necessity for permission of distribution. Social media sites such as Facebook, LinkedIn and WhatsApp, as well as third party participants who were reachable via email were also utilised. As previously stated, this type of sampling is known as convenience sampling in order to generate responses which can have a negative effect on the generalisability of the results (Etikan, 2016). To mitigate against this, the participants within the convenience sample were asked to share the survey link to their own friends, family and professional colleagues where possible. This is a term known as snowballing and it was employed to increase the number of respondents who
completed the survey and distribute it to a wider population (Dragan & Isaic-Maniu, 2013). This also allowed for greater generalisability of the results across a larger population set. Before the survey was distributed to the population, piloting was performed on five participants in order to confirm that the survey was understandable. All participants were informed of the format of the survey once they opened it, which clearly listed the instructions and what information would be collected. The survey did not obtain any identifiable information such as name or email address and each question was mandatory, to reduce the number of incorrect and thus unusable responses. Due to the data being divided among the various demographic groups, this could potentially harm the statistical legitimacy of the results. For example, if only four respondents are male and 50 are female, no statistical analysis can be concluded due to the imbalance of sample sizes (Button et al., 2013a). With this in mind, the target sample size was between eighty and one hundred participants, which should allow enough responses in order to accurately draw conclusions. The timeframe for the data collection phase was three weeks, to allow enough time to gather an adequate number of responses.

3.3.4 Preamble
When participants open the survey, they are first presented with a preamble as seen below, intended to provide context about the survey and how it should be undertaken. This was also done to reassure them of the anonymity of their responses, which was an important step due to the volume of demographic information collected.

This anonymous quiz contains 15 questions relating to three phishing categories (5 per category).

Section 1 will collect some demographic information then the questions are presented in Section 2.

Each question will contain an image and you will answer if the content is:

1. Real (you would interact with it like normal and follow the links presented)
   or

2. Fake (a malicious attempt to trick the user or steal information)

Please note, once you fill in this survey, your responses will be treated in a highly confidential manner. No identifiable information such as your name or email address is collected and no third party will have access to this data.
3.4 Statistical Tools and Methods Used

3.4.1 Variables
The survey collected data through variables known as categorical variables. Categorical variables provide a means for the information to be sorted into categories, for example, gender is either “Male” or “Female”. Every piece of information gathered belongs to one category and a respondent cannot be part of more than one category per demographic group. There are several types of categorical variables which the data belongs to. Ordinal variables relate to when the variables have some order applied to them. An example of this is recording income levels for a group of people in a variable called income. One person makes less than €20,000 per year, another makes between €20,000-40,000 while another makes €40,000 or more. These income levels fit nicely into defined categories as the income figure increases. From within the collected dataset, age, education and income are marked as ordinal variables.

Nominal variables are the opposite of ordinal variables because they do not contain any defined order. For example, a variable called location, which records where a person lives, is nominal because each of the categories do not have an order or numeric value. Due to this, gender and occupation are counted as nominal variables within the collected data set.

3.4.2 Two-sample T Test
A two-sample t test was used in order to analyse gender with the four test score dependant variables. Skaik (2015), states, “It is used to know whether the unknown means of two populations are different from each other based on independent samples from each population.” This test requires that the two independent variables are unrelated to each other and is only valid for comparing two means from a quantitative variable. The data comes from a single population which has been divided into two groups, male and female so it meets all the requirements in order to perform the two sample t test. Skaik (2015) goes on to say “If the study aims to compare three or more means, then it is better to use an analysis of variance to avoid the loss of control over the experiment-wise significant level.” A one-way analysis of variance (ANOVA) is used for the remaining four demographic groups as they contain more than two categories each.
3.4.3 One-way Anova

For the statistical analysis, a one-way analysis of variance (ANOVA) was selected for demographic groups that contained more than two categories. It is used to determine if the mean of a dependant variable is the same in three or more mutually independent variables. For this study, the dependant variables are:

1. Total Test Score
2. Category 1 Test Score
3. Category 2 Test Score
4. Category 3 Test Score

The independent variables refer to the different categories in each demographic group e.g. for education, they are:

1. Level 5 (Advanced/Higher Certificate)
2. Level 6 (Bachelor’s Degree/Higher Diploma)
3. Level 7 (Master's Degree/Postgraduate Diploma)
4. Level 8 (Doctorate Degree/Higher Doctorate)
5. None of the above

The one-way Anova uses the mean value from the dependant and independent variables to show if there is a difference between the two. This model was used to determine whether the differences between the categories were statistically significant, with the significance level or alpha set to 5% or ($\alpha = .05$). Lavrakas (2008), states, “Alpha is a threshold value used to judge whether a test statistic is statistically significant”. The author goes on to say, “Alpha represents an acceptable probability of a Type I error in a statistical test”. When a statistical significance is found between categories of a demographic group, it is possible to determine which category or categories were significantly different from each other by using a post hoc test. Post hoc tests are needed “because the one-way ANOVA is an omnibus test and cannot tell you which specific groups were significantly different from each other; it only tells you that at least two groups were different.” To alleviate against this, a post hoc test is run which is explained by Kim (2017). “When comparing the population means of three mutually independent groups A, B, and C, if the significance level is 0.05, then the significance level used for

---

comparisons of groups A and B, groups A and C, and groups B and C would be 0.05/3 = 0.017”.

Before a one way Anova can be run, there are six assumptions that need to be met in order to analyse the data. If any of the six assumptions are not met, it will not be possible to analyse the data effectively due to the results being invalid. The six assumptions of a one way Anova are:

**Assumption 1**: The dependant variable should be measured at the continuous level. The dependant variables of this study are test scores, which are continuously measured at either 0 to 5 or 0 to 15.

**Assumption 2**: “The independent variable should consist of two or more categorical, independent groups”. This assumption is met as each demographic group contains two or more unrelated groups.

**Assumption 3**: An independence of observations should be met so there is no relationship between the observations in each group or between groups. This assumption is also met.

**Assumption 4**: There should be no significant outliers within the dataset. These are simply values within the data that are much higher or lower than most other values, which could skew the results. This assumption is met because STATA can detect outliers when running the one way Anova.

**Assumption 5**: The “dependent variable should be approximately normally distributed for each category of the independent variable”. The one way Anova is quite robust to violations of these assumptions and due to this, the data needs to be approximately normal in order to provide valid results. The normality of the data was tested using the Shapiro-Wilk test in STATA as well as analysing the histogram which can be seen in Figure 3-1 below. “The word ‘normal’ here means that the data complies with a distribution pattern that mathematically allows parametric statistical tests to be applied” (Marshall & Jonker, 2010).

---

Assumption 6: Finally, using the Levene’s test in STATA, there needs to be homogeneity of variances. This was undertaken using the robvar model of Levene’s test for each variable and it was concluded there was equality of variances between the groups.

3.4.4 Statistical Software
In order to perform statistical analysis on the data collected from the survey, STATA was selected. STATA is “a complete, integrated statistical software package that provides everything you need for data analysis, data management, and graphics”.

3.5 Conclusion
This chapter outlined the design and methodology used in order to carry out the study. A detailed explanation of how the demographics are broken down was introduced such as using the European Qualifications Framework (EQF) for education. The reasons why a survey was selected for collecting data were addressed, stating that a convenience sample was the best method given the time constraints. The additional methods of piloting and using snowball sampling to increase the number of respondents was also

---

8 https://www.stata.com/why-use-stata/
addressed. A detailed look at the design of the study was given, breaking the survey
down into two parts and mentioning who was targeted for the respondent population.
Finally, a detailed explanation of how the data was to be captured, coded and analysed
was introduced. A two-sample t test and one-way Anova were selected as the statistical
tools to carry out the analysis.
4 Implementation and Results

4.1 Introduction

The object of this survey was to answer the research question, “Are members of demographic groups more vulnerable to being exploited by a particular category of phishing attack?” The survey collected demographic information and then presented the respondents with 15 questions relating to three separate phishing categories. This chapter gives a breakdown of the demographic distribution, as well as the noteworthy descriptive statistics that were gathered. Finally, the survey results are statistically analysed using STATA and it is concluded if there is enough evidence to reject the null hypotheses H0.

4.2 Survey Responses

4.2.1 Completion Rates

In total, 189 responses were collected from the survey. This is double the original target sample size, which will give more statistical power to the results. Each question in the survey was mandatory which meant that no incomplete answers could be submitted. Due to this, the number of unfinished surveys was not tracked.

4.2.2 Demographic Distribution

The first section of the survey collected demographic information of the respondents. A breakdown of the participant’s gender can be seen in figure 4-1 below. As demonstrated in the graph, there is an evenly balanced sample of respondents from each gender, which is pivotal for the study. Gender is one of the most commonly analysed demographics when it comes to phishing vulnerability. Due to this, an even distribution will aid the statistical analysis and allow the results to be accurately compared against previous studies.
Figure 4-1: Percentile distribution of gender

A distribution of the age categories is outlined in figure 4-2. Similar to gender, age is one of the most commonly analysed demographics in phishing vulnerability studies. The chart shows a clear reduction in the number of responses as the age group increases. This may be due to the sampling method that was employed which targeted social media to acquire participants.

Figure 4-2: Percentile distribution of age groups

The results for Education as shown by figure 4-3 indicate a plurality of respondents have a Bachelor’s degree or higher diploma. Almost one third of the participants have a Master’s degree, while 24% have either none of the defined degrees or a Level 5 Associate degree. This is valuable information to receive as analysis can be performed.
to see if having a Level 5 or lower education is detrimental to a person’s vulnerability to phishing.

The annual gross income categories as seen in figure 4-4 show a relatively even distribution. Due to income being one of the least analysed demographics as outlined in Chapter 2.6.3, having a sufficient sample size among each of the categories was important. Eleven percent of respondents chose not to disclose their gross income. In the context of this study, this information is redundant because no clear results can be drawn given that any of the participants may have selected this option. Due to this, the “I don’t want to disclose” group results will be disregarded during analysis even if they are below our significance level.

Figure 4-3: Percentile distribution of education levels
Finally, the results of the participants occupation is highlighted in figure 4-5. Over sixty percent of those surveyed are employed within either the Business and Finance industry or the Information Technology industry. Only six percent of the population are students, which is lower than initially anticipated. As discussed in chapter 2, students are being targeted by phishing attacks on a more regular basis so a large number of responses in this category were desired so statistically significant conclusions to be made. This uneven distribution may make it difficult to rely on the results of the statistical test, which are addressed in more detail in Chapter 4.6.5.

Figure 4-5: Percentile distribution of sectoral occupation/students
4.3 Descriptive Statistics

The data presented in this chapter highlights the descriptive statistics of each demographic group cross tabulated against total test scores, and their individual category test scores.

- **Total Quiz Scores** refers to the participants total score for all fifteen questions.
- **Category 1 Scores** refers to the participants score for category 1 i.e. Questions 1-5 relating to email phishing.
- **Category 2 Scores** refers to the participants score for category 2 i.e. Questions 6-10 relating to smishing.
- **Category 3 Scores** refers to the participants score for category 3 i.e. Questions 11-15 relating to content injection.

Descriptive statistics are used to provide a headline analysis about the data collected. They allow initial observations to be made regarding the data, and a chance to see potential trends. Once the descriptive statistics have been analysed, inferential statistics can be applied in order to draw conclusions and project the findings onto a larger group.

The mean and sample standard deviation were analysed as they make up the fundamentals of data analysis. The mean was selected because the data is consistent which will give a reliable answer each time. It is important not to rely on one statistic to describe the data, as outliers and skewed distributions may affect the reliability of findings. The sample standard deviation was also chosen because it is a commonly used measure of variability. Again, if the data has a non-normal distribution or many outliers, just using the standard deviation will not display all the information that is needed. Marshall & Jonker (2010) states “the standard deviation allows for expressing variance using the same units as those used for the observations or measurements”.

Due to the number of demographic groups and dependant variables, and with brevity in mind, only the most significant statistics have been included in this chapter. The full list of descriptive statistics are outlined in Appendix 2.

4.3.1 Gender

**Gender – Total Scores**

Table 4-1 shows the descriptive statistics for gender cross tabulated with their total test scores. Males have a slightly higher mean value compared to females.
Table 4-1: Descriptive statistics of the total test scores for gender

Age – Total scores
Table 4-2 shows the descriptive statistics for age cross tabulated with their total test scores. The 56 + age group has the lowest mean score, while the 26-35 bracket appears to have done the best, given the comparison with their standard deviation.

Table 4-2: Descriptive statistics of the total test scores for age

Age – Category 3 scores
Test scores for the different age groups from category three are displayed in table 4-3 below. Once again, the 56 + age group have performed the worst when analysing their mean value but the 36 – 45 age group have a much lower standard deviation. This indicates that the values in the statistical data set are close to the mean of the data set.

Table 4-3: Descriptive statistics of the category 3 test scores for age

4.3.2 Education
Education – Total scores
Table 4-4 below shows the descriptive statistics for education compared with their total test scores. For the purposes of clarity, the “None of these” category within the Education group is referred to as “No Degree” in this chapter. It can be seen that the “No
"Degree" group has the lowest mean and standard deviation value, indicating a correlation between lower education and vulnerability to phishing.

<table>
<thead>
<tr>
<th>No Degree</th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
<th>Level 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>20</td>
<td>24</td>
<td>77</td>
<td>50</td>
</tr>
<tr>
<td>Mean</td>
<td>9.75</td>
<td>10.20</td>
<td>11.20</td>
<td>11.14</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>1.58</td>
<td>2.18</td>
<td>1.64</td>
<td>2.32</td>
</tr>
</tbody>
</table>

Table 4-4: Descriptive statistics of the total test scores for education

**Education – Category 1 scores**

Category one test scores for the education demographic are displayed in table 4-5. It can be seen that the “No Degree” and “Level 5” groups perform the worst on average.

<table>
<thead>
<tr>
<th>No Degree</th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
<th>Level 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>20</td>
<td>24</td>
<td>77</td>
<td>50</td>
</tr>
<tr>
<td>Mean</td>
<td>3.35</td>
<td>3.45</td>
<td>3.88</td>
<td>3.94</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>1.03</td>
<td>1.28</td>
<td>0.99</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table 4-5: Descriptive statistics for the category 1 test scores for education

**4.3.3 Income**

**Income – Total scores**

Table 4-6 below shows the descriptive statistics for income compared with their total test scores. There is a relatively large mean test score difference among the €20,000-30,000 group when compared with the others.

<table>
<thead>
<tr>
<th>€</th>
<th>0-20k</th>
<th>20-30k</th>
<th>30-40k</th>
<th>40-50k</th>
<th>50-60k</th>
<th>60k+</th>
<th>Not Disclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>18</td>
<td>30</td>
<td>43</td>
<td>29</td>
<td>22</td>
<td>26</td>
<td>21</td>
</tr>
<tr>
<td>Mean</td>
<td>11.05</td>
<td>10.40</td>
<td>10.97</td>
<td>11.06</td>
<td>10.86</td>
<td>11.07</td>
<td>11.19</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>3.11</td>
<td>2.35</td>
<td>1.48</td>
<td>1.53</td>
<td>1.67</td>
<td>1.74</td>
<td>1.47</td>
</tr>
</tbody>
</table>

Table 4-6: Descriptive statistics of the total test scores for the income demographic
Income – Category 1 scores

The category one test scores of the income demographic, the 60,000 + group scored best. This group’s standard deviation value when compared to the others, also seem to indicate a greater spread in the data as seen in Table 4-7.

<table>
<thead>
<tr>
<th>€</th>
<th>0-20k</th>
<th>20-30k</th>
<th>30-40k</th>
<th>40-50k</th>
<th>50-60k</th>
<th>60k+</th>
<th>Not Disclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong></td>
<td>18</td>
<td>30</td>
<td>43</td>
<td>29</td>
<td>22</td>
<td>26</td>
<td>21</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>3.77</td>
<td>3.43</td>
<td>3.86</td>
<td>3.95</td>
<td>3.59</td>
<td>4.00</td>
<td>3.76</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>1.24</td>
<td>1.25</td>
<td>0.90</td>
<td>0.86</td>
<td>0.95</td>
<td>1.01</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table 4-7: Descriptive statistics of the category 1 test scores for the income demographic

4.3.4 Occupation/Student

Occupation/Student – Total scores

Table 4-8 shows the descriptive statistics for occupation compared with their total test scores. The media and healthcare occupations have the best mean scores, with agriculture scoring the worst.

<table>
<thead>
<tr>
<th>Occupation/Student</th>
<th>Agri</th>
<th>Bus</th>
<th>Edu</th>
<th>Health</th>
<th>IT</th>
<th>Manu</th>
<th>Media</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong></td>
<td>4</td>
<td>73</td>
<td>14</td>
<td>17</td>
<td>51</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>9.0</td>
<td>10.56</td>
<td>10.28</td>
<td>11.23</td>
<td>11.05</td>
<td>10.62</td>
<td>11.60</td>
<td>11.33</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>2.94</td>
<td>1.73</td>
<td>1.63</td>
<td>4.69</td>
<td>1.30</td>
<td>0.74</td>
<td>1.57</td>
<td>1.92</td>
</tr>
</tbody>
</table>

Table 4-8: Descriptive statistics of the total test scores for the occupation/student demographic

Occupation/Student – Category 1 scores

Test scores for the occupation groups from category 1 are displayed in table 4-9. The mean scores are even across all occupations apart from agriculture which performed the worst.

<table>
<thead>
<tr>
<th>Occupation/Student</th>
<th>Agri</th>
<th>Bus</th>
<th>Edu</th>
<th>Health</th>
<th>IT</th>
<th>Manu</th>
<th>Media</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong></td>
<td>4</td>
<td>73</td>
<td>14</td>
<td>17</td>
<td>51</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>2.75</td>
<td>3.76</td>
<td>3.71</td>
<td>3.88</td>
<td>3.88</td>
<td>3.87</td>
<td>3.50</td>
<td>3.83</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>2.06</td>
<td>0.99</td>
<td>0.99</td>
<td>1.05</td>
<td>0.90</td>
<td>0.99</td>
<td>0.97</td>
<td>1.26</td>
</tr>
</tbody>
</table>

Table 4-9: Descriptive statistics of the category 1 test scores for the occupation/student demographic
Occupation/Student – Category 3 scores

Test scores for the different occupation groups from category 3 are displayed in table 4-10.

<table>
<thead>
<tr>
<th></th>
<th>Agri</th>
<th>Bus</th>
<th>Edu</th>
<th>Health</th>
<th>IT</th>
<th>Manu</th>
<th>Media</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>4</td>
<td>73</td>
<td>14</td>
<td>17</td>
<td>51</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Mean</td>
<td>2</td>
<td>2.46</td>
<td>2.07</td>
<td>2.76</td>
<td>3.17</td>
<td>2.50</td>
<td>3.40</td>
<td>2.83</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>0</td>
<td>0.92</td>
<td>0.82</td>
<td>0.97</td>
<td>0.99</td>
<td>0.75</td>
<td>0.96</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Table 4-10: Descriptive statistics of the category 3 test scores for the occupation/student demographic

4.3.5 Additional Analysis

Additional analysis was performed to identify the breakdown of correct answers per question. It is clear from figure 4-6 that question 11 and 15 were the most incorrectly answered questions by a large margin. Both phishing attempts required the user to analyse the URL of the webpage. Although it is a subtle difference, by doing so they would have realised it was a fake site. This confirms the points by Wu, Miller, & Garfinkel (2006), stating that “Users fail to continuously check the browser’s security indicators, since maintaining security is not the user’s primary goal”.

![Correct Answers per Question](image)

Figure 4-6: The percentage of correct answers per question in the survey

4.4 Data Clean up

Once the data collections were complete and the survey was closed, the data was manually analysed and cleaned. This was possible due to the relatively small response
size and completed in order for it to be used by a statistical program. The data was exported to a spreadsheet and the variables were coded so they could be imported into STATA. The variables before being coded were qualitative i.e. Gender consisted of “Male” and “Female”. In order to facilitate analysis they were given numerical values starting with zero and incrementing through the different demographic categories. This is done so statistical models can be run across the data, so STATA can understand the variables, and models can be run “blind” in order to reduce bias. A breakdown of the categories after being coded can be seen in table 4-11 below.

<table>
<thead>
<tr>
<th>Value</th>
<th>Gender</th>
<th>Age</th>
<th>Education</th>
<th>Income</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Female</td>
<td>18-25</td>
<td>None of these</td>
<td>None of these</td>
<td>Agriculture</td>
</tr>
<tr>
<td>1</td>
<td>Male</td>
<td>26-35</td>
<td>Level 5</td>
<td>0 - 20,000</td>
<td>Bus/Finance</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>36-45</td>
<td>Level 6</td>
<td>20,000 - 30,000</td>
<td>Education</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>46-55</td>
<td>Level 7</td>
<td>30,000 - 40,000</td>
<td>Healthcare</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>56 +</td>
<td>Level 8</td>
<td>40,000 - 50,000</td>
<td>IT</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>50,000 - 60,000</td>
<td>Manufacturing</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>60,000 +</td>
<td>Media</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Student</td>
</tr>
</tbody>
</table>

Table 4-11: A breakdown of the demographic categories with coding applied

4.5 Theoretical Approach

In inferential statistics, the term ‘null hypothesis’ (\( H_0 \), ‘\( H-naught \), ‘\( H-null \)’) denotes that there is no relationship (difference) between the population variables in question. Alternative hypothesis (\( H_1 \) and \( H_a \)) denotes that a statement between the variables is expected to be true. (Kim, 2017)

The findings in this chapter are displayed using the following statistical information as seen in tables 4-12 and 4-13 below. For hypothesis 1, a two-sample t-test is used which will analyse the 2-tailed p-value (\( \Pr(|T| > |t|) \)) and also the mean value (contrast) and standard error.
T-Test

| 2-tailed p-value $Pr(|T| > |t|)$ (p) | Mean (contrast) ± standard error |
|-----------------------------------|---------------------------------|

Table 4-12: Two sample t-test values, which are to be analysed

The formula for the t-test can be seen in figure 4-7.

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Figure 4-7: T-test formula

For the remaining hypotheses, a one-way Anova will be run initially and the results will be presented using the F-value and the 2-tailed p-value $[Prob > F]$. As mentioned in Chapter 3.4.3, if the p value is below the significance level ($\alpha = .05$), a post hoc test will be run to get a better understanding of the groups within the demographic set. If a post hoc test is run across the data, it is analysed using the mean (contrast) ± standard error and the 2-tailed p-value $[Prob > |t|]$ to identify any statistically significant differences between the categories.

One Way Anova | Post Hoc Test
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>F-value (F)</td>
<td>Mean (contrast) ± standard error</td>
</tr>
<tr>
<td>2-tailed p-value $[Prob &gt; F]$ (p)</td>
<td>2-tailed p-value $[Prob &gt;</td>
</tr>
</tbody>
</table>

Table 4-13: Values for the one-way Anova and post hoc test which are to be analysed

4.6 Empirical Results

The empirical results are presented in this chapter using the statistical models highlighted in Chapter 4.5. The inclusion of asterisks beside values are used to denote different significance levels.
At the end of each hypothesis sub chapter, there is a breakdown of the four category test scores. They highlight the p value, whether the results are statistically significant and thus, if we reject the null hypothesis.

4.6.1 Hypothesis 1
As specified in Chapter 3.4.2, a two-sample t test was chosen in order the measure phishing vulnerability between genders. This statistical test uses the mean value of the four dependant variables with the independent variable, gender. The two-sample t test was run first using the total quiz scores dependant variable, which was the participant’s total score out of all fifteen questions. This was used to give a broader picture of the participant’s vulnerability to phishing, before analysing the individual categories as well. Participants were broken down by males and females (n = number of participants):

- Male (n = 92)
- Female (n = 97)

The first hypothesis of this study is stated below:

**Hypothesis 1:**

**H₀:** When given a quiz, there is no difference in the mean test scores between males and females as to who are more vulnerable to fall for a phishing attack.

**H₁:** When given a quiz, there is a difference in the mean test scores between males and females as to who are more vulnerable to fall for a phishing attack.

**Total Test Scores:**
For total test scores, there was not a statistically significant difference when comparing males and females. The results show that females did not have a statistically significant vulnerability to being phished (10.77 ± .1868) when compared with males (11.09 ± .1665), p = 0.1978. This p value is outside of the defined significance level (α = .05). With these results, the null hypothesis H₀ will not be rejected.
Category 1 and 2 Test Scores:
Similar to total test scores, category one and two test scores did not return a statistically significant difference when comparing males and females. The results show that females did not have a statistically significant vulnerability to being phished via email (3.86 ± .1011) when compared with males (3.68 ± .1081), p = 0.2220. The same result can be seen when analysing female’s vulnerability to smishing (4.39 ± .07967) compared with males (4.47 ± .6644), p = 0.4080. With these results, the null hypothesis H₀ will not be rejected for both categories.

Category 3 Test Scores:
There was a statistically significant difference between the category three test scores when comparing gender vulnerability. The results show that females had a statistically significant vulnerability to being phished via content spoofing (2.51 ± .0939) when compared with males (2.93 ± .1090), p = 0.0039***. Due to this, the null hypothesis H₀ can be rejected and the alternative hypothesis H₁ can be accepted.

The p value for each test score and whether the null hypothesis is rejected can be seen in table 4-14.

<table>
<thead>
<tr>
<th>Gender – Analysis</th>
<th>P Value</th>
<th>Statistically Significant</th>
<th>Reject H₀</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Test Scores</td>
<td>0.1978</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Category 1 Scores</td>
<td>0.2220</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Category 2 Scores</td>
<td>0.4080</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Category 3 Scores</td>
<td>0.0039***</td>
<td>Yes</td>
<td>Yes (Accept H₁)</td>
</tr>
</tbody>
</table>

Table 4-14: Analysis of results among gender regarding the four test scores

4.6.2 Hypothesis 2
For the remainder of the hypotheses, a one-way Anova was conducted to determine vulnerability to phishing among the demographics. Similar to the first hypothesis, the total quiz scores dependant variable was analysed first. This was used to give a broader picture of the participant’s vulnerability to phishing as a whole, before analysing the individual test score categories. Participants were classified into five age groups as previously mentioned (n = number of participants):
The second hypothesis of this study is highlighted below:

**Hypothesis 2:**

\( H_0 \): When given a quiz, there is no difference in the mean test scores between age groups as to who are more vulnerable to fall for a phishing attack.

\( H_1 \): When given a quiz, there is a difference in the mean test scores between age groups as to who are more vulnerable to fall for a phishing attack.

**Total Test Scores:**

There was a very statistically significant difference between the age categories \((p = 0.0002)\). Due to the \( p \) value registering below the chosen significance value of .05, a post hoc test was run to get a better understanding of how the specific groups differed. A Tukey post hoc test was used combined with a Pairwise comparison of means with equal variances. The Tukey post hoc test revealed that vulnerability to phishing was significantly higher in the 56+ age group when compared with several of the other age groups as highlighted below in Figure 4-15.

<table>
<thead>
<tr>
<th>Age Group Comparison</th>
<th>( P ) Value</th>
<th><strong>Contrast ± Std. Err.</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>56+ vs 18-25</td>
<td>0.045**</td>
<td>-1.42 ± .509</td>
</tr>
<tr>
<td>56+ vs 26-35</td>
<td>0.001***</td>
<td>-3.88 ± .510</td>
</tr>
<tr>
<td>46-55 vs 26-35</td>
<td>0.002***</td>
<td>-1.38 ± .366</td>
</tr>
</tbody>
</table>

Table 4-15: Statistically significant results for the age demographic when analysing their total test scores

Due to these results, the null hypothesis \( H_0 \) can be rejected and the alternative hypothesis \( H_1 \) can be accepted.
**Category 1 Test Score:**
Similar to the total test scores, there was a statistically significant difference between the age groups (F = 3.82, p = 0.0052***). A Tukey post hoc test was run in order to get a better understanding of the differences between the age groups. Only one group was statistically significant from the other, with the 46-55 group being more vulnerable when compared against the 36-45 age group as (-.677 ± .235, p = 0.036**). Following on from the total test scores, the null hypothesis H₀ will be rejected and alternative hypothesis H₁ will be accepted.

**Category 2 Test Scores:**
When analysing the category two test scores for age, there is a statistically significant difference between the age groups (F = 3.80, p = 0.0054***). The post hoc test revealed that vulnerability to smishing was significantly higher in the 46-55 age group compared to the 18-25 age group (-.550 ± .154, p = 0.004***). There was no statistically significant difference between any of the other groups. Given these results, the null hypothesis H₀ can be rejected and the alternative hypothesis H₁ can be accepted.

**Category 3 Test Scores:**
For the category three test scores, there is a statistically significant difference between the age groups (F = 4.80, p = 0.0010***). The post hoc test revealed that vulnerability to content spoofing was significantly higher in the 36-45 and 56+ age groups compared to the 26-35 age group (-.773 ± .206, p = 0.002*** (-.961 ± .299, p = 0.14)) respectively. However, there was no statistically significant difference between any of the other groups. Due to these results, the null hypothesis H₀ can be rejected and the alternative hypothesis H₁ can be accepted. The p value for each test score and whether the null hypothesis is rejected can be seen in table 4-16.
<table>
<thead>
<tr>
<th>Age – Analysis</th>
<th>P Value</th>
<th>Statistically Significant</th>
<th>Reject H0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Test Scores</td>
<td>0.0002***</td>
<td>Yes</td>
<td>Yes (Accept H1)</td>
</tr>
<tr>
<td>Category 1 Scores</td>
<td>0.0052***</td>
<td>Yes</td>
<td>Yes (Accept H1)</td>
</tr>
<tr>
<td>Category 2 Scores</td>
<td>0.0054***</td>
<td>Yes</td>
<td>Yes (Accept H1)</td>
</tr>
<tr>
<td>Category 3 Scores</td>
<td>0.0010***</td>
<td>Yes</td>
<td>Yes (Accept H1)</td>
</tr>
</tbody>
</table>

Table 4-16: Analysis of results among age regarding the four test scores

4.6.3 Hypothesis 3

A one-way Anova was conducted to determine if the education level had an effect on a user’s vulnerability to phishing. Participants were classified into five groups as previously mentioned (n = number of participants). For the purposes of clarity, the “None of these” group is referred to as “No Degree” in this chapter.

- Level 5 (n = 24)
- Level 6 (n = 77)
- Level 7 (n = 50)
- Level 8 (n = 18)
- No Degree (n = 20)

The third hypothesis of this study is highlighted below:

**Hypothesis 3:**

H<sub>0</sub>: When given a quiz, there is no difference in the mean test scores between education levels as to who are more vulnerable to fall for a phishing attack

H<sub>1</sub>: When given a quiz, there is a difference in the mean test scores between education levels as to who are more vulnerable to fall for a phishing attack

**Total Test Scores:**

There was a statistically significant difference between the total test score when compared with education levels (F = 4.81, p = 0.0010***). The Tukey post hoc test revealed that vulnerability to phishing was significantly higher in the No Degree group when compared with Level 6, Level 7 and Level 8 education as seen below.
Table 4-17: Statistically significant results for the education demographic when analysing their total test scores

This indicates that the No Degree group has a negative 1.69 lower mean score than Level 8 and this result is found to be statistically significant at the .05 level with (p = 0.017). It can be seen that the contrast value increases with the level of education when compared with the no degree group. This would indicate that the more educated the user is, the less vulnerable they are to being phished. Due to these results, the null hypothesis H₀ can be rejected and the alternative hypothesis H₁ can be accepted.

**Category 1 Test Scores:**

For category one test scores, there was not a statistically significant difference when comparing them with education levels (F = 2.04, p = 0.0901). No post hoc test was run due to the p value registering higher than the significance level of .05. With these results, the null hypothesis H₀ will not be rejected.

**Category 2 Test Scores:**

There was not a statistically significant difference when comparing the category two test scores with education levels (F = 2.35, p = 0.0561). No post hoc test was run due to the p value registering higher than the significance level of .05. With these results, the null hypothesis H₀ will not be rejected.

**Category 3 Test Scores:**

There was a statistically significant difference between the category three test scores when compared with education levels (F = 2.43, p = 0.0490*). A post hoc test was run but no results were within the 0.05 significance level. Due to this, the null hypothesis H₀ can be rejected and the alternative hypothesis H₁ can be accepted.

The p value for each test score and whether the null hypothesis is rejected can be seen in table 4-18.
**Education Analysis** | **P Value** | **Statistically Significant** | **Reject H₀**
--- | --- | --- | ---
**Total Test Scores** | 0.0010*** | Yes | Yes (Accept H₁)
**Category 1 Scores** | 0.0901 | No | No
**Category 2 Scores** | 0.0561 | No | No
**Category 3 Scores** | 0.0490* | Yes | Yes (Accept H₁)

Table 4-18: Analysis of results among education regarding the four test scores

### 4.6.4 Hypothesis 4

A one-way Anova was conducted to determine if income level had an effect on a user’s vulnerability to phishing. Participants were classified into seven income brackets as previously mentioned (n = number of participants).

- None of these (n = 21)
- 0 - 20,000 (n = 18)
- 20,000 - 30,000 (n = 30)
- 30,000 - 40,000 (n = 43)
- 40,000 - 50,000 (n = 29)
- 50,000 - 60,000 (n = 22)
- 60,000 + (n = 26)

The fourth hypothesis of this study is highlighted below:

**Hypothesis 4:**

**H₀:** When given a quiz, there is no difference in the mean test scores between income levels as to who are more vulnerable to fall for a phishing attack

**H₄:** When given a quiz, there is a difference in the mean test scores between income levels as to who are more vulnerable to fall for a phishing attack

### Total Test Scores

There was not a statistically significant difference between the total test score when compared with income levels (F = 0.63, p = 0.7057). Due to the p value registering above the significance value of .05, a post hoc test was not run against the data. From the results, the null hypothesis H₀ will not be rejected.
Category 1, 2 and 3 Test Scores:
Similar to the total test scores, there was not a statistically significant difference between either the category one, two or three test scores when compared with income levels as seen in table 4-19 below. A post hoc test was not run and because of the results, the null hypothesis $H_0$ will not be rejected.

<table>
<thead>
<tr>
<th>Income Analysis</th>
<th>$F$ Value</th>
<th>$P &gt; F$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cat 1 Scores</td>
<td>1.12</td>
<td>0.3522</td>
</tr>
<tr>
<td>Cat 2 Scores</td>
<td>1.70</td>
<td>0.1229</td>
</tr>
<tr>
<td>Cat 3 Scores</td>
<td>0.31</td>
<td>0.9335</td>
</tr>
</tbody>
</table>

Table 4-19: Category 1, 2 and 3 scores when running a one-way Anova on gross income brackets

An overall analysis of the results from the income demographic can be seen in table 4-20.

<table>
<thead>
<tr>
<th>Income – Analysis</th>
<th>$P$ Value</th>
<th>Statistically Significant</th>
<th>Reject $H_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Test Scores</td>
<td>0.7057</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Category 1 Scores</td>
<td>0.3522</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Category 2 Scores</td>
<td>0.1229</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Category 3 Scores</td>
<td>0.9335</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 4-20: Analysis of results among income regarding the four test scores

4.6.5 Hypothesis 5
A one-way Anova was conducted to determine if occupation or being a student had an effect on a user’s vulnerability to phishing. Participants were classified into eight fields as previously mentioned (n = number of participants).

- Agriculture (n = 4)
- Business/Finance (n = 73)
- Education (n = 14)
- Healthcare (n = 17)
- Information Technology (n = 51)
- Manufacturing (n = 8)
• Media (n = 10)
• Student (n = 12)

The large distributional difference and varying sample sizes between the occupation demographics as seen above raises issues when employing a statistical model. Various studies highlight the importance of an evenly distributed sample size in order to properly infer results (Button et al., 2013). Button et al., (2013), states, “Low statistical power undermines the purpose of scientific research; it reduces the chance of detecting a true effect”. Statistical power is affected primarily by the size of the sample used to detect it\(^9\). Due to this, it will be difficult to rely on these results and generalise them onto a larger population.

The fifth hypothesis of this study is highlighted below:

**Hypothesis 5:**

**H\(_0\):** When given a quiz, there is no difference in the mean test scores between occupations or students as to who are more vulnerable to fall for a phishing attack.

**H\(_5\):** When given a quiz, there is a difference in the mean test scores between occupations or students as to who are more vulnerable to fall for a phishing attack.

**Total Test Scores**

There was a statistically significant difference between the total test score when compared with occupation (\(F = 2.89, p = 0.0069^{***}\)). With the p value registering below the significance value of .05, a post hoc test was run to get a better understanding of how the specific groups differed. The Tukey post hoc test revealed that vulnerability to phishing was significantly higher in the Business and Finance occupation when compared with Information Technology (\.9481 \pm .305, p = 0.044^*\). There were no other statistically significant results returned from the post hoc test. From these results, the null hypothesis \(H_0\) will be rejected and the alternative hypothesis \(H_5\) will be accepted.

**Category 1 Test Scores:**

For category one test scores, there was not a statistically significant difference when comparing different occupations (\(F = 0.81, p = 0.5794\)). No post hoc test was run due to

\(^9\) https://effectsizefaq.com/2010/05/31/what-is-statistical-power/
the p value registering higher than the significance level of .05. With these results, the null hypothesis $H_0$ will not be rejected.

**Category 2 Test Scores:**
There was not a statistically significant difference when comparing category two test scores with the different occupations ($F = 0.85, p = 0.5513$). No post hoc test was run due to the p value registering higher than the significance level of .05. With these results, the null hypothesis $H_0$ will not be rejected.

**Category 3 Test Scores:**
There was a statistically significant difference between the category three test scores when compared with occupation ($F = 4.56, p = 0.0001^{***}$). The post hoc test revealed that the Business and Finance occupation is more vulnerable to content spoofing when compared with Information Technology professionals as seen in table 4-21. This coincides with the results from the total test scores. It can also be seen that Education professionals were more vulnerable to content spoofing when compared with both Information Technology and Media professionals. From these results, the null hypothesis $H_0$ will be rejected and the alternative hypothesis $H_1$ will be accepted.

<table>
<thead>
<tr>
<th>Occupation Cat 3 Statistics</th>
<th>P Value</th>
<th>Contrast ± Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT vs Bus/Finance</td>
<td>0.001***</td>
<td>.710 ± .172</td>
</tr>
<tr>
<td>IT vs Education</td>
<td>0.004***</td>
<td>1.105 ± .285</td>
</tr>
<tr>
<td>Media vs Education</td>
<td>0.019**</td>
<td>1.328 ± .391</td>
</tr>
</tbody>
</table>

Table 4-21: Statistically significant results for the occupation demographic when analysing their category 3 test scores

The results obtained from the occupation demographic are highlighted in Table 4-22 below.
<table>
<thead>
<tr>
<th>Occupation Analysis</th>
<th>P Value</th>
<th>Statistically Significant</th>
<th>Reject H₀</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Test Scores</td>
<td>0.0069***</td>
<td>Yes</td>
<td>Yes (Accept H₁)</td>
</tr>
<tr>
<td>Category 1 Scores</td>
<td>0.5794</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Category 2 Scores</td>
<td>0.5513</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Category 3 Scores</td>
<td>0.0001***</td>
<td>Yes</td>
<td>Yes (Accept H₁)</td>
</tr>
</tbody>
</table>

Table 4-22: Analysis of results among occupation/students regarding the four test scores

4.7 Conclusion

This chapter outlined the implementation and results from the statistical analysis. A breakdown of the demographic distribution was highlighted, showing how respondents were spread between the different groups. The noteworthy descriptive statistics were presented, which allowed initial observations to be made before the statistical models were run across the data. The method for how the data was cleaned and coded so it could be imported into STATA was also introduced. Finally, the theoretical approach was undertaken, running the two-sample t-test and one-way Anova models across the results of the survey in order to see if there was a statistically significant outcome. From these results, each of the five hypotheses were answered.
5 Analysis, Evaluation and Discussion

5.1 Introduction
This chapter will discuss the results obtained in chapter four compared with what is already known. Whether the results of this project coincide or refute previous studies will be addressed and reasons for these results are discussed.

5.1.1 Gender
Previous research has found mixed results when analysing phishing vulnerability among gender (Kumaraguru et al., 2009a; Sheng et al., 2010). Some studies concluded that there was no significant difference among gender in the tendency to fall for phishing emails (Kumaraguru et al., 2009a), while others established women as the more vulnerable gender (Sheng et al., 2010). The findings from this study illustrate that there is no statistically significant difference between males and females when comparing their vulnerability to phishing as a whole or through email and smishing. This indicates that neither males nor females are more vulnerable to phishing when analysing the above three categories. However, when considering the category 3 scores, women appear to be more susceptible to content spoofing when compared to men. These results would suggest that females are more vulnerable than males at falling for content spoofing phishing attacks. Perhaps this could be contributed to women’s tendency to be more trustworthy of web content, as suggested in a study by Gavett et al., (2017).

5.1.2 Age
All of the previous research undertaken appears to highlight that younger adults, mainly in the 18-25 age group are the most vulnerable to phishing attacks (Gavett et al., 2017; Kumaraguru et al., 2009a; Sheng et al., 2010). From the results obtained in Chapter 4.6.2, it was concluded that the 56 + age group was the most vulnerable to being phished when analysing participants total test scores. This can be seen not only when examining the mean and standard deviation values but also when using a post hoc test to compare vulnerability between all age groups.

Following on from this, a consistent pattern emerges when analysing the category one, two and three test scores of each age group. It was observed that the 46-55 age group was more vulnerable to email phishing when compared with the 36-45 group for category one test scores. For category two scores, the 46-55 age group was again more vulnerable to being caught out via smishing when compared against the 18-25 age group.
Finally, when analysing category three test scores, vulnerability to content spoofing was significantly higher in the 36-45 and 56+ age groups compared to the 26-35 age group. All of these results appear to indicate that older adults above the age of 36 are more vulnerable than younger adults to being phished in all categories that were analysed. This goes against the findings from previous work already obtained and may be down to a number of factors. Potential explanations for this include that older individuals are not as well informed or experienced with technology. Younger adults who grow up with a multitude of technology at their fingertips are more likely to be able to identify a spoofed phishing request through experience and knowledge. This same concept can be thought for smishing vulnerability as many of the younger participants may be more experienced with their mobile phones and can identify the phishing attempts easier.

5.1.3 Education

In the limited research that has been carried out to analyse education levels and vulnerability to phishing, Dhamija et al., (2006), concluded that there was “no significant correlation between education level and scores”. However, when analysis is performed on the results of this survey, a different outcome is presented. For total test scores, there was a very statistically significant difference when comparing education levels. From the results obtained, participants with no degree i.e. those who selected “none of these” within the survey were most vulnerable to phishing. In fact, there was a direct correlation between low-test scores and no degree when analysing the mean and standard deviation as well as running the one-way Anova. Furthermore, vulnerability to phishing was significantly higher in the no degree group when compared with those who have Level 6, Level 7 and Level 8 education.

There are a number of potential explanations for these results. Perhaps participants who do not have a degree are in the first years of third level education. Those who have completed higher-level education potentially have access to greater information technology and education resources. There has been no other research into the influence of education on vulnerability to phishing. This would suggest a gap in the literature and understanding into the relationship between this demographic, and suggests the need for further research.
5.1.4 Income

Previous research suggest that there is no evidence linking a high income with an increased vulnerability to being phished (Leukfeldt 2014). The findings from this study coincide with previous research, as there were no statistically significant differences between income levels and test scores.

Given these results, phishing is portrayed in a slightly darker, more ruthless light because it is deemed that no one is safe. Leukfeldt (2014) stated “An unemployed person on a shoestring budget or a director of a multinational company: everyone has an equal chance of becoming a victim.”

5.1.5 Occupation/Student

From the previous research that was reviewed, it revealed that students were more prone to phishing attacks when compared against working professionals (Metzger, Flanagin, & Zwarun 2003; Mohebzada et al., 2012). Mohebzada et al., (2012) concluded, “Experience could be a factor in victimising an individual during a social engineering attack”. The results from the analysis go against these findings, as there is no statistically significant difference between students and any of the occupations when comparing their vulnerability to phishing across all categories. Each of the student participants were identified as being in the 18-25 age group. With this information, a correlation could be suggested with the age demographic results in chapter 5.1.2. The results found younger participants less likely to be phished when compared with older adults, which could be a result of information technology experience and knowledge.

It was difficult to locate any peer-reviewed literature that compared and contrasted different occupations and their vulnerability to phishing. Due to this, there are no studies that can be referenced against the results of this project. For total test scores and category three test scores, there was a statistically significant difference when comparing test scores and occupation. It transpired that Business and Finance professionals could be considered as more vulnerable to being phished when compared with Information Technology professionals. These findings suggest that IT professionals are more aware of attempted phishing attacks which could be attributed to industry knowledge and experience. For category three results, respondents who are in education were also seen as more vulnerable to phishing attacks when compared with IT professionals. Education professionals were also seen as more vulnerable, when compared with Media professionals, to content spoofing attacks.
5.2 Conclusion

This chapter reviewed the results of the analysis obtained in chapter four and identified whether they agreed or contested findings of previous studies. Finally, a discussion on why these results were obtained was introduced.
6 Conclusion and Future Work

6.1 Introduction
This chapter reviews the objectives and limitations of the research, the contributions to the body of knowledge and any recommendations for future work that may be undertaken.

6.2 Research Overview and Problem Definition
The objective of this research project was to answer the following research question:

*Are members of demographic groups more vulnerable to being exploited by a particular category of phishing attack?*

The emphasis of the study was focused on three types of phishing, which are currently viewed as some of the most prevalent and dangerous. Measuring a user’s vulnerability to a specific type of phishing is an important aspect in order to help mitigate against its threat. If it is clear who is most vulnerable to a specific type of phishing attack, it is possible that more secure methods of anti-phishing training and awareness could be introduced to reduce the quantity of successful attacks.

Quantitative research was performed in order to statistically analyse if particular members of a demographic group were more vulnerable to a specific type of phishing. The primary data gathered from the survey was analysed using statistical tools in order to identify if a statistically significant result was obtained.

6.3 Limitations of Research
Undertaking a study of this size can introduce limitations that may potentially halt or hinder the research.

- One of the main limitations experienced during the course of this study was a lack of previous research when analysing demographic vulnerability to phishing. Although there were a handful of studies that could be used to aid the research, the small amount of information affected the discussion when considering results against the backdrop of previous literature.

- A limitation encountered when performing the analysis was a lack of cross analysis between demographics. The statistical analysis performed looked at singular demographics and the members within them but no analysis was
undertaken to consider how multiple demographics may affect vulnerability to phishing.

- A true random sample would have been helped to increase the likelihood that the results will accurately reflect the entire population (Brant, Haas-Haseman, Wei, Wickham, & Ponto 2015).

6.4 Contributions and Impact

This dissertation focused on identifying differences that exist among five demographics to determine which members are the most vulnerable to being exploited by a particular type of phishing. As previously mentioned in Chapter 1, no study has been undertaken to determine what types of phishing are the most successful at exploiting users. The results from this study were able to identify which members of a demographic group were vulnerable to a specific kind of phishing.

As previously mentioned, the lack of literature to examine demographic vulnerability has limited this study when discussing the results. Therefore, the findings from this study are significantly relevant to examining which members of a demographic group are vulnerable to phishing. This study is expected to contribute the results, conclusions and recommendations to the information security body of knowledge. It is hoped that this material will aid future researchers as well as contribute to other studies in this field.

6.5 Future Work & Recommendations

There are gaps in the literature to suggest that future work is needed in order to continuously monitor user’s vulnerability to new and improved phishing attacks.

- One area where future research regarding demographic phishing vulnerability can benefit from is cross analysis. The use of regression analysis between demographics can help to determine what characteristics make users vulnerable. In the context of this study, it could be used to determine what combination of demographics make a user the most vulnerable e.g. 56 + year old males who works in the Business and Finance industry.

- In light of these results, another potential area of future work is to use these findings in order to develop innovative ways to help stop phishing attacks. This
study highlights who is at risk but a deeper study is required in order to leverage these findings and come up with innovative ways to mitigate phishing attacks.

6.6 Final Thoughts

The results from this study clearly highlight that phishing is still as deceitful and dangerous as ever; with the average total test score being 72%. Although this may seem like a good score in academic terms, it only takes one phishing fraud to potentially cause massive harm to a victim both financially and emotionally. It is important that studies like these continue to be undertaken so consistent and up to date information can be gathered. The results obtained from additional studies can be used to inform anti-phishing work groups, educate people about the dangers of phishing and help mitigate against future attacks.
Appendices

Appendix A - Survey Questions

Section 1 – Demographic Information

Gender

☐ Male

☐ Female

Age *

☐ 18-25

☐ 26-35

☐ 36-45

☐ 46-55

☐ > 56

Education (EFQ Framework) *

☐ Level 5 (Advanced/Higher Certificate)

☐ Level 6 (Bachelors Degree/Higher Diploma)

☐ Level 7 (Master’s Degree/Postgraduate Diploma)

☐ Level 8 (Doctorate Degree/Higher Doctorate)

☐ None of the above
**Income** *

- €0 - 20,000
- €20,000 - 30,000
- €30,000 - 40,000
- €40,000 - 50,000
- €50,000 - 60,000
- €60,000 +
- I don't want to disclose

**Student or Working? Please list Industry** *

- Student
- Information Technology
- Education
- Business/Finance
- Manufacturing
- Healthcare
- Agriculture
- Media
Section 2 – Phishing Questions

Question 1

Important: We noticed unusual activity in your PayPal account

What's going on?

We're concerned that someone is using your PayPal account without your knowledge. Recent activity on your account seems to have occurred from a suspicious location or under circumstances that may be different than usual.

What to do?

Log in to your PayPal account as soon as possible. We may ask you to confirm information you provided when you created your account to make sure you're the account holder. We'll then ask you to confirm your password and security questions. You should also do the following for your own protection:

Confirm Your Account Now

Log in to confirm your account

Question 2

DO (via LinkedIn)
Mon 27/11/2017, 07:21

invited you to connect 9 days ago.

Accept View Invitation

Global Head of Sales & Marketing at
Machines Ltd.

View profile
Dear Customer,

Your ID was just used to purchase "Fire TV Print HD 89.97 €" from the Store on a computer or device that had not previously been associated with that ID.

This purchase was initiated from Mexico.

If you made this purchase, you can disregard this email. It was only sent to alert you in case you did not make the purchase yourself.

If you did not make this purchase, we recommend you to click Here to update your information immediately.

---

*Amazon.co.uk* is a trading name for Amazon EU S.à.r.l. (Luxembourg Registration Number B-101818, VAT Number LU 20289743) and Amazon Media EU S.à.r.l. (Luxembourg Registration Number 112767, VAT Number LU 39944528). Each company is located at 5 Rue Adolphe, L-3338 Luxembourg.

Please note that this promotional e-mail is being sent from an e-mail address that cannot receive e-mails. If you have any questions and wish to contact us, click here.
Dear Default,

Each month we automatically attempt to renew your subscription based on the preferences you selected upon signing up with Netflix. We have recently attempted to debit your registered payment method and unfortunately the payment attempt has failed.

We have temporarily limited your services until your account has been validated.

Please verify that the information below is correct:

Name: Default  
Payment Type: VISA  
Card Number: XXXX-XXXX-XXXX  
Expiry Date: [redacted]

Once you have confirmed that the above information is correct, please click the button below to validate your account in order for your subscription to continue.

Validate my Account
Question 5

New device signed in to

Your Google Account was just signed in to from a new Windows device. You’re getting this email to make sure it was you.

CHECK ACTIVITY

Question 6

Messages to this chat and calls are now secured with end-to-end encryption. Tap for more info.

ALDI is giving away £150 Free Voucher to celebrate 27th anniversary, go here to get yours: http://www.aldi.com/celebration.
Question 7 *

Dear NAB Bank User,
We have detected some unusual activity.
We urgently ask you to follow the account review link:
http://bit.do/nab-bank

Question 8 *

Dear Customer,

Your AppleID is due to expire Today, Please tap http://bit.do/cRqb6 to update and prevent loss of services and data.

Apple  smsSTOPto43420
Question 9

Hello! Your PayPal account was compromised on September 13th at 08:00 UTC time. Please sign in and confirm the authenticity of your recent transactions. http://paypayl.com/unlock account/

Question 10

<From 3: Your new ebill is now ready to view. For the full bill go to My3 online or on your mobile. For a summary click on http://mobile.3ireland.ie/sms04 >
Question 11

[Image of Outlook Web App login page]

Question 12

[Image of Google Doc login page]

Access Shared Documents with any Valid Email Address

You can now access Google with all valid email accounts:

- Yahoo!
- Hotmail
- AOL
- GMail
- Others
Question 13 *
Question 14.
Appendix B - Descriptive Statistics

Gender

Gender – Total Scores
Table B-1 shows the descriptive statistics for gender cross tabulated with their total test scores.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Quiz Scores</td>
<td>1021</td>
<td>1045</td>
</tr>
<tr>
<td>Participants</td>
<td>92</td>
<td>97</td>
</tr>
<tr>
<td>Mean</td>
<td>11.09</td>
<td>10.77</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>1.59</td>
<td>1.83</td>
</tr>
</tbody>
</table>

Table B-1: Descriptive statistics of the total test scores for gender

Gender – Category 1 Scores
Table B-2 shows the descriptive statistics for gender cross tabulated with their category one test scores.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 Scores</td>
<td>339</td>
<td>375</td>
</tr>
<tr>
<td>Participants</td>
<td>92</td>
<td>97</td>
</tr>
<tr>
<td>Mean</td>
<td>3.68</td>
<td>3.86</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>1.03</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table B-2: Descriptive statistics of the category 1 test scores for gender

Gender – Category 2 Scores
Table B-3 shows the descriptive statistics for gender cross tabulated with their category two test scores.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 2 Scores</td>
<td>412</td>
<td>426</td>
</tr>
<tr>
<td>Participants</td>
<td>92</td>
<td>97</td>
</tr>
<tr>
<td>Mean</td>
<td>4.47</td>
<td>4.39</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>0.63</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Table B-3: Descriptive statistics of the category 2 test scores for gender
Gender – Category 3 Scores

Table B-4 shows the descriptive statistics for gender cross tabulated with their category three test scores.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category 3 Scores</strong></td>
<td>270</td>
<td>244</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td>92</td>
<td>97</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>2.93</td>
<td>2.51</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>1.04</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table B-4: Descriptive statistics of the category 3 test scores for genders

Age

Age – Total scores

Table B-5 shows the descriptive statistics for age cross tabulated with their total test scores.

<table>
<thead>
<tr>
<th>Age – Total scores</th>
<th>18 - 25</th>
<th>26 - 35</th>
<th>36 - 45</th>
<th>46 - 55</th>
<th>56 +</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Quiz Scores</strong></td>
<td>585</td>
<td>603</td>
<td>416</td>
<td>337</td>
<td>125</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td>53</td>
<td>52</td>
<td>38</td>
<td>33</td>
<td>13</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>11.03</td>
<td>11.59</td>
<td>10.94</td>
<td>10.21</td>
<td>9.61</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>1.77</td>
<td>1.50</td>
<td>1.55</td>
<td>1.62</td>
<td>1.59</td>
</tr>
</tbody>
</table>

Table B-5: Descriptive statistics of the total test scores for age

Age – Category 1 scores

Table B-6 shows the descriptive statistics for age cross tabulated with their category 1 test scores.

<table>
<thead>
<tr>
<th>Age – Category 1 scores</th>
<th>18 - 25</th>
<th>26 - 35</th>
<th>36 - 45</th>
<th>46 - 55</th>
<th>56 +</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category 1 Scores</strong></td>
<td>192</td>
<td>208</td>
<td>157</td>
<td>114</td>
<td>43</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td>53</td>
<td>52</td>
<td>38</td>
<td>33</td>
<td>13</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>3.62</td>
<td>4</td>
<td>4.13</td>
<td>3.45</td>
<td>3.30</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>1.10</td>
<td>0.89</td>
<td>0.86</td>
<td>1.01</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Table B-6: Descriptive statistics of the category 1 test scores for age
Age – Category 2 scores

Table B-7 shows the descriptive statistics for age cross tabulated with their category 2 test scores.

<table>
<thead>
<tr>
<th></th>
<th>18 - 25</th>
<th>26 - 35</th>
<th>36 - 45</th>
<th>46 - 55</th>
<th>56 +</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category 2 Scores Participants</strong></td>
<td>246</td>
<td>233</td>
<td>170</td>
<td>135</td>
<td>54</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td>53</td>
<td>52</td>
<td>38</td>
<td>33</td>
<td>13</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>4.64</td>
<td>4.48</td>
<td>4.47</td>
<td>4.09</td>
<td>4.15</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>0.55</td>
<td>0.69</td>
<td>0.59</td>
<td>0.89</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Table B-7: Descriptive statistics of the category 2 test scores for age

Age – Category 3 scores

Table B-7 shows the descriptive statistics for age cross tabulated with their category 3 test scores.

<table>
<thead>
<tr>
<th></th>
<th>18 - 25</th>
<th>26 - 35</th>
<th>36 - 45</th>
<th>46 - 55</th>
<th>56 +</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category 3 Scores Participants</strong></td>
<td>147</td>
<td>162</td>
<td>89</td>
<td>88</td>
<td>28</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td>53</td>
<td>52</td>
<td>38</td>
<td>33</td>
<td>13</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>2.77</td>
<td>3.11</td>
<td>2.3</td>
<td>2.66</td>
<td>2.15</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>1.05</td>
<td>1.01</td>
<td>0.13</td>
<td>0.87</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Table B-8: Descriptive statistics of the category 3 test scores for age
Education

Education – Total scores

Table B-9 shows the descriptive statistics for education cross tabulated with their total test scores.

<table>
<thead>
<tr>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
<th>Level 8</th>
<th>No Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Quiz Scores</td>
<td>245</td>
<td>863</td>
<td>557</td>
<td>206</td>
</tr>
<tr>
<td>Participants</td>
<td>24</td>
<td>77</td>
<td>50</td>
<td>18</td>
</tr>
<tr>
<td>Mean</td>
<td>10.20</td>
<td>11.20</td>
<td>11.14</td>
<td>11.44</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>2.18</td>
<td>1.64</td>
<td>2.32</td>
<td>1.90</td>
</tr>
</tbody>
</table>

Table B-9: Descriptive statistics of the total test scores for education

Education – Category 1 scores

Table B-10 shows the descriptive statistics for education cross tabulated with their category 1 test scores.

<table>
<thead>
<tr>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
<th>Level 8</th>
<th>No Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 Scores</td>
<td>83</td>
<td>299</td>
<td>197</td>
<td>68</td>
</tr>
<tr>
<td>Participants</td>
<td>24</td>
<td>77</td>
<td>50</td>
<td>18</td>
</tr>
<tr>
<td>Mean</td>
<td>3.45</td>
<td>3.88</td>
<td>3.94</td>
<td>3.77</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>1.28</td>
<td>0.99</td>
<td>0.91</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Table B-10: Descriptive statistics of the category 1 test scores for education

Education – Category 2 scores

Table B-11 shows the descriptive statistics for education cross tabulated with their category 2 test scores.

<table>
<thead>
<tr>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
<th>Level 8</th>
<th>No Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 2 Scores</td>
<td>107</td>
<td>345</td>
<td>220</td>
<td>85</td>
</tr>
<tr>
<td>Participants</td>
<td>24</td>
<td>77</td>
<td>50</td>
<td>18</td>
</tr>
<tr>
<td>Mean</td>
<td>4.45</td>
<td>4.48</td>
<td>4.4</td>
<td>4.72</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>0.5</td>
<td>0.6</td>
<td>0.728</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Table B-11: Descriptive statistics of the category 2 test scores for education
Education – Category 3 scores

Table B-12 shows the descriptive statistics for education cross tabulated with their category 3 test scores.

<table>
<thead>
<tr>
<th></th>
<th>Level 5</th>
<th>Level 6</th>
<th>Level 7</th>
<th>Level 8</th>
<th>No Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category 3 Scores</strong></td>
<td>55</td>
<td>219</td>
<td>140</td>
<td>53</td>
<td>47</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td>24</td>
<td>77</td>
<td>50</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>2.29</td>
<td>2.84</td>
<td>2.8</td>
<td>2.94</td>
<td>2.35</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>1.04</td>
<td>0.87</td>
<td>1.06</td>
<td>1.05</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Table B-12: Descriptive statistics of the category 3 test scores for education

Income

Income – Total scores

Table B-13 shows the descriptive statistics for income cross tabulated with their total test scores.

<table>
<thead>
<tr>
<th></th>
<th>0-20k</th>
<th>20-30k</th>
<th>30-40k</th>
<th>40-50k</th>
<th>50-60k</th>
<th>60k+</th>
<th>Not Disclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Quiz Scores</strong></td>
<td>199</td>
<td>312</td>
<td>472</td>
<td>321</td>
<td>239</td>
<td>288</td>
<td>235</td>
</tr>
<tr>
<td><strong>Participants</strong></td>
<td>18</td>
<td>30</td>
<td>43</td>
<td>29</td>
<td>22</td>
<td>26</td>
<td>21</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>11.05</td>
<td>10.4</td>
<td>10.97</td>
<td>11.89</td>
<td>10.86</td>
<td>11.076</td>
<td>11.190</td>
</tr>
<tr>
<td><strong>Sample SD, s</strong></td>
<td>3.11</td>
<td>2.35</td>
<td>1.48</td>
<td>1.53</td>
<td>1.67</td>
<td>1.74</td>
<td>1.47</td>
</tr>
</tbody>
</table>

Table B-13: Descriptive statistics of the total test scores for income
Income – Category 1 scores

Table B-14 shows the descriptive statistics for income cross tabulated with their category 1 test scores.

<table>
<thead>
<tr>
<th>Income – Category 1 scores</th>
<th>0-20k</th>
<th>20-30k</th>
<th>30-40k</th>
<th>40-50k</th>
<th>50-60k</th>
<th>60k+</th>
<th>Not Disclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>18</td>
<td>30</td>
<td>43</td>
<td>29</td>
<td>22</td>
<td>26</td>
<td>21</td>
</tr>
<tr>
<td>Mean</td>
<td>3.77</td>
<td>3.43</td>
<td>3.86</td>
<td>3.96</td>
<td>3.59</td>
<td>4</td>
<td>3.761</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>1.24</td>
<td>1.25</td>
<td>0.9</td>
<td>0.86</td>
<td>0.95</td>
<td>1.01</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table B-14: Descriptive statistics of the category 1 test scores for income

Income – Category 2 scores

Table B-15 shows the descriptive statistics for income cross tabulated with their category 2 test scores.

<table>
<thead>
<tr>
<th>Income – Category 2 scores</th>
<th>0-20k</th>
<th>20-30k</th>
<th>30-40k</th>
<th>40-50k</th>
<th>50-60k</th>
<th>60k+</th>
<th>Not Disclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>18</td>
<td>30</td>
<td>43</td>
<td>29</td>
<td>22</td>
<td>26</td>
<td>21</td>
</tr>
<tr>
<td>Mean</td>
<td>4.77</td>
<td>4.333</td>
<td>4.302</td>
<td>4.44</td>
<td>4.5</td>
<td>4.26</td>
<td>4.66</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>0.42</td>
<td>0.42</td>
<td>0.74</td>
<td>0.63</td>
<td>0.80</td>
<td>0.82</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Table B-15: Descriptive statistics of the category 2 test scores for income
Income – Category 3 scores

Table B-16 shows the descriptive statistics for income cross tabulated with their category 3 test scores.

<table>
<thead>
<tr>
<th>Category 3 Scores</th>
<th>0-20k</th>
<th>20-30k</th>
<th>30-40k</th>
<th>40-50k</th>
<th>50-60k</th>
<th>60k+</th>
<th>Not Disclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>18</td>
<td>30</td>
<td>43</td>
<td>29</td>
<td>22</td>
<td>26</td>
<td>21</td>
</tr>
<tr>
<td>Mean</td>
<td>2.5</td>
<td>2.63</td>
<td>2.81</td>
<td>2.65</td>
<td>2.77</td>
<td>2.80</td>
<td>2.7</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>0.92</td>
<td>1.15</td>
<td>0.90</td>
<td>0.81</td>
<td>1.03</td>
<td>1.05</td>
<td>1.26</td>
</tr>
</tbody>
</table>

Table B-16: Descriptive statistics of the category 3 test scores for income

Occupation/Student

Occupation/Student – Total scores

Table B-17 shows the descriptive statistics for occupation cross tabulated with their total test scores.

<table>
<thead>
<tr>
<th>Agri</th>
<th>Bus</th>
<th>Edu</th>
<th>Health</th>
<th>IT</th>
<th>Manu</th>
<th>Media</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Quiz Scores</td>
<td>36</td>
<td>771</td>
<td>144</td>
<td>191</td>
<td>587</td>
<td>85</td>
<td>116</td>
</tr>
<tr>
<td>Participants</td>
<td>4</td>
<td>73</td>
<td>14</td>
<td>17</td>
<td>51</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Mean</td>
<td>9</td>
<td>10.56</td>
<td>10.28</td>
<td>11.23</td>
<td>11.0</td>
<td>10.6</td>
<td>11.6</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>2.94</td>
<td>1.73</td>
<td>1.63</td>
<td>4.69</td>
<td>1.30</td>
<td>0.74</td>
<td>1.57</td>
</tr>
</tbody>
</table>

Table B-17: Descriptive statistics of the total test scores for occupation

Occupation/Student – Category 1 scores

Table B-18 shows the descriptive statistics for occupation cross tabulated with their category 1 test scores.

<table>
<thead>
<tr>
<th>Agri</th>
<th>Bus</th>
<th>Edu</th>
<th>Health</th>
<th>IT</th>
<th>Manu</th>
<th>Media</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 Scores</td>
<td>11</td>
<td>275</td>
<td>52</td>
<td>66</td>
<td>198</td>
<td>31</td>
<td>35</td>
</tr>
<tr>
<td>Participants</td>
<td>4</td>
<td>73</td>
<td>14</td>
<td>17</td>
<td>51</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Mean</td>
<td>2.75</td>
<td>3.76</td>
<td>3.71</td>
<td>3.88</td>
<td>3.88</td>
<td>3.87</td>
<td>3.5</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>2.06</td>
<td>0.99</td>
<td>0.99</td>
<td>1.05</td>
<td>0.90</td>
<td>0.99</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table B-18: Descriptive statistics of the category 1 test scores for occupation
Occupation/Student – Category 2 scores

Table B-19 shows the descriptive statistics for occupation cross tabulated with their category 2 test scores.

<table>
<thead>
<tr>
<th></th>
<th>Agri</th>
<th>Bus</th>
<th>Edu</th>
<th>Health</th>
<th>IT</th>
<th>Manu</th>
<th>Media</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 Scores</td>
<td>17</td>
<td>316</td>
<td>63</td>
<td>78</td>
<td>227</td>
<td>34</td>
<td>47</td>
<td>56</td>
</tr>
<tr>
<td>Participants</td>
<td>4</td>
<td>73</td>
<td>14</td>
<td>17</td>
<td>51</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Mean</td>
<td>4.0</td>
<td>4.32</td>
<td>4.5</td>
<td>4.58</td>
<td>4.45</td>
<td>4.25</td>
<td>4.7</td>
<td>4.66</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>0.95</td>
<td>0.8</td>
<td>0.65</td>
<td>0.79</td>
<td>0.6</td>
<td>0.70</td>
<td>0.48</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Table B-19: Descriptive statistics of the category 2 test scores for occupation

Occupation/Student – Category 3 scores

Table B-20 shows the descriptive statistics for occupation cross tabulated with their category 3 test scores.

<table>
<thead>
<tr>
<th></th>
<th>Agri</th>
<th>Bus</th>
<th>Edu</th>
<th>Health</th>
<th>IT</th>
<th>Manu</th>
<th>Media</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 Scores</td>
<td>8</td>
<td>180</td>
<td>29</td>
<td>47</td>
<td>162</td>
<td>20</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>Participants</td>
<td>4</td>
<td>73</td>
<td>14</td>
<td>17</td>
<td>51</td>
<td>8</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Mean</td>
<td>2</td>
<td>2.46</td>
<td>2.07</td>
<td>2.76</td>
<td>3.17</td>
<td>2.5</td>
<td>3.4</td>
<td>2.83</td>
</tr>
<tr>
<td>Sample SD, s</td>
<td>0</td>
<td>0.92</td>
<td>0.82</td>
<td>0.97</td>
<td>0.99</td>
<td>0.75</td>
<td>0.96</td>
<td>1.11</td>
</tr>
</tbody>
</table>

Table B-20: Descriptive statistics of the category 3 test scores for occupation
7 Bibliography


83


