

Spring 3-23-2021

Analyzing the Effectiveness of Legal Regulations and Social Consequences for Securing Data

Howard B. Goodman
Dakota State Universi

Follow this and additional works at: <https://scholar.dsu.edu/theses>



Part of the [Information Security Commons](#), [OS and Networks Commons](#), and the [Other Computer Sciences Commons](#)

Recommended Citation

Goodman, Howard B., "Analyzing the Effectiveness of Legal Regulations and Social Consequences for Securing Data" (2021). *Masters Theses & Doctoral Dissertations*. 370.
<https://scholar.dsu.edu/theses/370>

This Dissertation is brought to you for free and open access by Beadle Scholar. It has been accepted for inclusion in Masters Theses & Doctoral Dissertations by an authorized administrator of Beadle Scholar. For more information, please contact repository@dsu.edu.



Analyzing the Effectiveness of Legal Regulations and Social Consequences for Securing Data

A dissertation submitted to the Beacom College of Computer and Cyber Sciences at

Dakota State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Cyber Operations

June 21, 2021

By

Howard B. Goodman

Dissertation Committee:

Dr. Pam Rowland

Dr. Shengjie Xu

Dr. Chris Olson

Mary Francis

DISSERTATION APPROVAL FORM

This dissertation is approved as a credible and independent investigation by a candidate for the Doctor of Philosophy degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this dissertation does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department or university.

Student Name: Howard B. Goodman

Dissertation Title: Analyzing the Effectiveness of Legal Regulations and Social Consequences for Securing Data

Dissertation Chair: Pam Rowland Date: March 31, 2021
Dr. Pam Rowland

Committee member: Shengjie Xu Date: March 31, 2021
Dr. Shengjie Xu

Committee member: Chris Olson Date: March 31, 2021
Dr. Chris Olson

Committee member: Mary Francis Date: March 31, 2021
Mary Francis

ACKNOWLEDGMENTS

From the years I've spent at Dakota State University, I have come to understand the level of commitment it takes to complete a doctorate. I am deeply grateful to so many people - faculty, staff, fellow students, friends, and family who contributed to my development as a scholar and researcher. My dissertation is but one manifestation of the trials over the past four years.

I would like to offer a special thanks to my Chair, Dr. Pam Rowland; without your encouragement and counsel, I would never have known where to start. Also, words cannot express my gratitude to Dr. Shengjie Xu, Dr. Chris Olson, and Mary Francis for challenging me and for all their commitment and support. Thank you all for enthusiastically advising me on ways to make my research stronger. Finally, I would like to thank everyone that took the time to participate in my survey. Without their responses, I would have never had the data I need to complete my research.

ABSTRACT

There is a wide range of concerns and challenges related to stored data security – which range from privacy and management to operations readiness, These challenges span from financial to personal and public impact. With an abundance of regulations for the enforcement of data security and emerging requirements proposed every year, organizations cannot avoid the legal or social implications of inadequate data protection. Today, public spotlight and awareness are challenging organizations to enhance how data is protected more than at any other time. For this reason, organizations have made significant efforts to improve security.

When looking at precautions or changes, the factors considered are costs associated with such action, a potential consequence of not acting, impact on users, the effort required, and the scope. For this reason, leaders need to make the hard decisions of which risks they can live with and which need to be reduced because it is unrealistic to think that data security can be guaranteed. However, it is essential to have physical, administrative, and technical controls to mitigate data risks. Data protection regulations define requirements, create procedures to identify the associated risks, determine the extent of the impact, and identify what precautions should be taken.

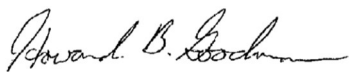
This dissertation defined seven areas for consideration related to stored data security. The research facilitated developing a measurement tool to gather and analyze the knowledge and opinions of working professionals within the United States. The study was performed from July to October 2020, which resulted in a quantitative data sample used to analyze the effectiveness of legal regulations and social consequences for securing data.

DECLARATION

I hereby certify that this project constitutes my own product, that where the language of others is set forth, quotation marks so indicate, and that appropriate credit is given where I have used the language, ideas, expressions, or writings of another.

I declare that the project describes original work that has not previously been presented for the award of any other degree of any institution.

Signed,



Howard B. Goodman

TABLE OF CONTENTS

Dissertation Approval Form	ii
Acknowledgments.....	iii
Abstract	iv
Declaration	v
List of Figures	ix
List of Tables	xi
Chapter 1: Introduction	1
1.1 Background	2
1.2 Statement of Problem.....	4
1.3 Research Questions and Hypotheses.....	5
1.4 Objective and Contribution to the Discipline.....	6
1.5 Limitation.....	6
1.6 Dissertation Structure.....	6
Chapter 2: Literature Review.....	8
2.1 Prevailing Definitions	8
2.2 Article Analysis	11
2.3 Source Classification.....	14
2.4 Article Mapping to CIA Triad.....	14
2.5 Results of Classification.....	16
2.6 Principles of stored data security	18
2.6.1 Confidentiality.....	19
2.6.2 Integrity	21
2.6.3 Availability.....	24
2.7 Security Controls	27
2.8 Regulatory Data Protection Models	29
2.8.1 Ensuring compliance	30
2.8.2 Comprehensive Data Protection Model.....	31

2.8.3	Sectoral Data Protection Model	32
2.8.4	Co-Regulatory and Self-Regulatory Models	32
2.8.5	Selected Regulations	33
2.9	Consequences and Penalties.....	35
2.10	Summary.....	38
Chapter 3: Research Methodology.....		38
3.1	Population and Sample Size Requirements.....	39
3.2	Scope and Approach	41
3.3	Research Tool Development and Approval	42
3.4	Demographic Categorization	43
3.5	Response Validation	44
3.6	Data Collection	45
3.7	Data Processing Analysis.....	46
3.8	Data Analysis Plan.....	51
3.8.1	Descriptive Statistics	52
3.8.2	Binary Logistic Regression	52
3.8.3	Point Biserial Correlation.....	53
3.8.4	Spearman Correlation.....	53
3.8.5	Two-Tailed Independent Samples t-Test.....	54
3.8.6	Data sources	54
3.9	Summary.....	55
Chapter 4: Results and Analysis		56
4.1	Demographics	56
4.2	Data Types	59
4.3	Stored Data Security Principles Scoring Results	60
4.4	Regulations	61
4.5	Research Question 1 Analysis.....	62
4.5.1	Descriptive Analysis	63
4.5.2	Binary Logistic Regression	64
4.5.3	Point Biserial Correlation Analysis.....	66
4.5.4	Two-Tailed Independent Samples t-Test.....	69

4.6	Research Question 2 and 3 Analysis	72
4.6.1	Descriptive Analysis	73
4.6.2	Spearman Correlation Analysis.....	74
4.7	Summary.....	79
Chapter 5: Conclusion.....		80
5.1	Limitations	80
5.2	Discussion of Findings.....	81
5.2.1	Research Question 1	82
5.2.2	Research Question 2 and 3	85
5.3	Summary and Future Works	89
References.....		93
Appendix A: Regulations Studied		117
Appendix B: Survey.....		128
Appendix C: IRB Approvals.....		135
Appendix D: Industry Categories		136
Appendix E: organization Size		142
Appendix F: Role Category		144
Appendix G: Survey Questions Classification with Reference number		148
Appendix H: Questionnaire Scoring Rubric		153
Appendix I: Glossary of Statistical Terms.....		158
Appendix J: Binary Logistic Regression Report		161
Appendix K: Point Biserial Correlation Report.....		167
Appendix L: Independent Samples t-Test report.....		171
Appendix M: Spearman Correlation Analysis Report 1		181
Appendix N: Spearman Correlation Analysis Report 2.....		184

LIST OF FIGURES

Figure 1 Leading Problems for Data Storage Worldwide in 2016 and 2017 (Liu, 2017)	2
Figure 2 IT Assets Vulnerable to Insider Threats (ENISA, 2019)	3
Figure 3 Data Types Vulnerable to Insider Threats (ENISA, 2019)	4
Figure 4 Security Principles Shown by Academic and Industry Articles.....	17
Figure 5 Security Principles Compared by Data and Storage Articles	18
Figure 6 RPO, RTO, MTD, and WRT (Marek.Z, 2013)	24
Figure 7 Cochran’s formula for calculating sample size (Cochran, 1977).....	40
Figure 8 Distribution of qualified and disqualified responses.....	45
Figure 9 Authentication Scoring probability distribution passing score > 2	50
Figure 10 Scoring curves for other security principles.....	51
Figure 11 Questions that derive the primary independent and dependent variables	55
Figure 12 Number of qualified participants based on role category and organization size.....	57
Figure 13 Number data types selected by qualified participants	59
Figure 14 Pass/fail sums by Security Principle.....	60
Figure 15 Total number of passed principles (<i>TotalPrinciplesPassed</i>).....	61
Figure 16 Regulation’s choices selected.....	62
Figure 17 Regulations where <i>TotalPrinciplesPassed</i> < 7	64
Figure 18 Present age of all regulations failing principle vs. None and Unknown combined	68
Figure 19 The mean of <i>TotalPrinciplesPassed</i> comparing regulations selected vs. not selected .	71
Figure 20 Scatterplots between each consequence with the regression line added	76
Figure 21 The difference for <i>TotalPrinciplesPassed</i> mean values regulation variables	85
Figure 22 Ranking of consequences vs. Stored Data Security	86

Figure 23 Spearman's ranking for each consequence variable and *SecPriVal* 87

Figure 24 Ranking of between *ConseqFines* and *SecPriVal* for *TotalPrinciplesPassed* 88

LIST OF TABLES

Table 1 Top Article Sources	12
Table 2 Years Published Grouped in three-year increments	12
Table 3 Principles cataloged under CIA Triad	13
Table 4 Confidentiality	14
Table 5 Integrity	15
Table 6 Availability	15
Table 7 Source Type by Concentration of Security Interest.....	16
Table 8 Seven Principles cross-referenced to Security Control classifications.....	28
Table 9 Regulations mapped to the seven principles of stored data security	35
Table 10 U.S. cost for data breaches 2019 – 2020 (Ponemon Institute & IBM Security, 2020)..	37
Table 11 Common Z-Table for standard Confidence Levels (LTCC, 2009).....	40
Table 12 Average time spent on the survey.....	45
Table 13 Results of validation check	46
Table 14 Summary Statistics Table for Interval and Ratio Variables (Intellectus Statistics, 2020)	48
Table 15 Calculating passing value criteria > 64% and < 70%	49
Table 16 Qualified responses by Industry split by Role (green = highest %)	57
Table 17 Industry split by organization size (Green = highest).....	58
Table 18 Industry split by top 8 selected data types (blue = Top 3).....	59
Table 19 percentages of failed principles by each regulation response.....	63
Table 20 Logistic Regression results with TotalPrinciplesPassed Approx. % change predicting the regulation	65
Table 21 Point Biserial correlations for Regulations and TotalPrinciplesPassed.....	67

Table 22 Two-Tailed Independent Samples t-Test for TotalPrinciplesPassed.....	69
Table 23 Averages of all means and total principles failed.....	71
Table 24 RQ2 and RQ3 variable names for analysis.....	73
Table 25 Consequences of incidence vs. importance of data security and privacy.....	73
Table 26 Spearman correlation results between consequences and SecPriVal.....	77
Table 27 Spearman correlation results between consequences and SecPriVal with TotalPrinciplesPassed.....	78
Table 28 Top 3 stored data security principles that failed for each selected regulation variable.	84
Table 29 Ranking of most likely failed stored data security principle.....	84

CHAPTER 1: INTRODUCTION

Modern data centers have accumulated a staggering volume of critical data for organizations, regardless of the industry vertical (IDG, 2017). Data has become the most crucial aspect of how an organization strategizes and justifies short- and long-term goals and actions across all operations. Storing data is fundamental to retaining information and requires the highest security controls, which is why data security requires meaningful, comprehensive, and simple strategies to mitigate threats.

As a result of a detailed literature review conducted in 2019 on data and storage security, a new framework model was created for seven key security areas that map to Confidentiality, Integrity, and Availability (CIA Triad). Chapter 2 goes into detail on the model. The research was peer-reviewed, presented at the National Cyber Summit 2020, and published in Springer's *Advances in Intelligent Systems and Computing* under the title "Deficiencies of Compliancy for Data and Storage Isolating the CIA Triad Components to Identify Gaps to Security" (Goodman & Rowland, 2020). The new model, subsequently referred to as the seven principles of stored data security, is used as the foundation of this dissertation research.

This dissertation examined the effectiveness of legal and social deterrents for securing stored data. It also used the seven principles of stored data security (or seven principles) to determine if regulations positively or negatively impact data security. In addition, it measured the effectiveness of various consequences to cyber and data incidence to determine which influenced an organization to implement stricter data security controls.

1.1 Background

From 2016 to 2017, 451 Research, a technology industry research firm acquired by S&P Global in 2019 (451 Group, 2020), conducted a study asking IT decision-makers about their main concerns regarding data storage. Figure 1 is an abbreviated version of 13 areas considered as the leading causes of stored data issues (Liu, 2017).

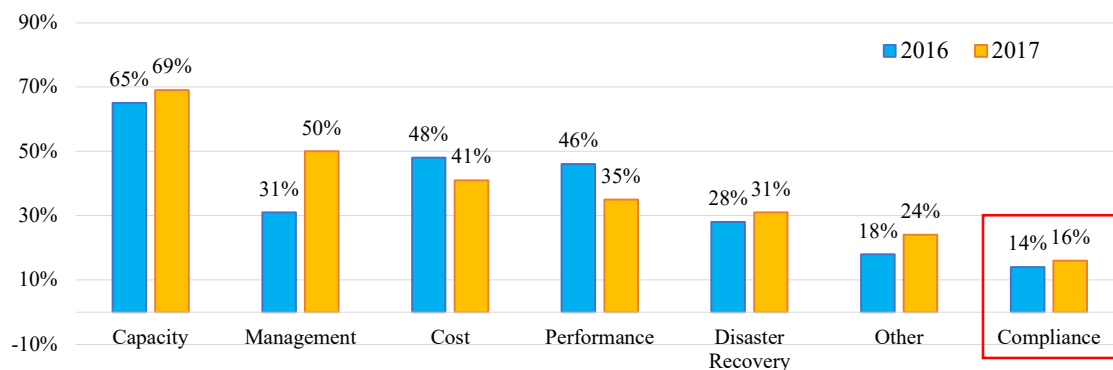


Figure 1 Leading Problems for Data Storage Worldwide in 2016 and 2017 (Liu, 2017)

The top four concerns found were: capacity, management, cost, and performance (Liu, 2017). The majority of the respondents did not see disaster recovery (DR) as a storage security concern, despite being the last resort for regaining regular operation after a cyber incident or natural disaster (Patterson, 2018). The research also discovered significant gaps showing that functional requirements are far more valuable than security concerns. Greater than 80% do not consider compliance a problem related to data storage (Liu, 2017).

While data storage has paved the way for dramatic advancements in the information age, it has become the highest-profile objective to threat actors. Threat actors or cyber threat actors are individuals, groups, or nation-states whose goal is to take advantage of vulnerabilities to gain

unauthorized access to systems, organizational IT environments, or secured data (Canadian Centre for Cyber Security, 2020). In the 2018 European Union Agency for Cybersecurity (ENISA) Threat Landscape Report 2018, the IT assets vulnerable to insider threats (see Figure 2) show that the top targets are containers for structured and unstructured data (ENISA, 2019).

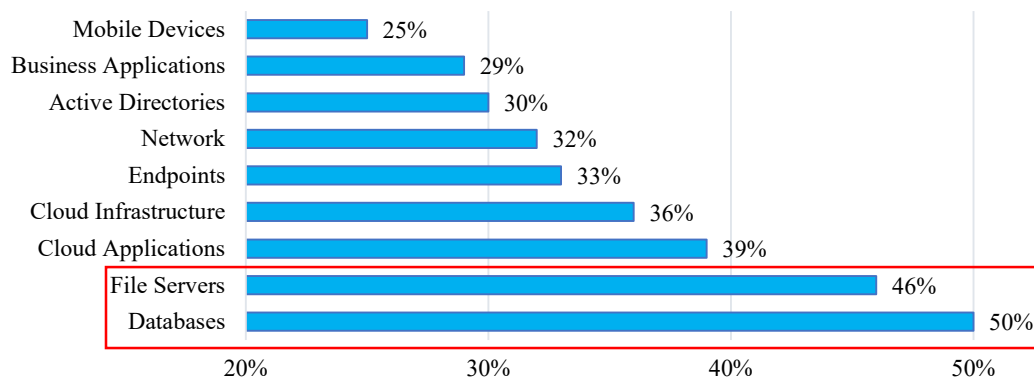


Figure 2 IT Assets Vulnerable to Insider Threats (ENISA, 2019)

While the various attack surfaces are the targets, the goal is to protect sensitive information. Digital information is so commonplace that most people hardly give it a second thought that public or private organizations store it when authorized (Herrera, 2019). However, organizations are aware of the data types necessary for routine operations. The same ENISA report is a breakdown of the data types identified as the most vulnerable (See Figure 3).

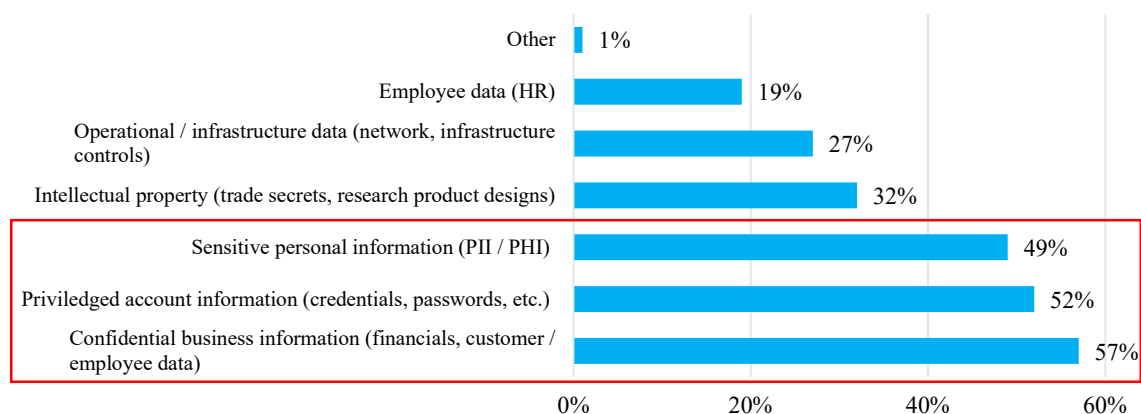


Figure 3 Data Types Vulnerable to Insider Threats (ENISA, 2019)

Data types related to confidential business information, user access credentials, Protected Health Information (PHI), and Personally Identifiable Information (PII) are at the most significant risk (ENISA, 2019). The most influential data security driver today is data privacy. As public awareness grows, so does the demand for more data security. While personal data security is driven by public demand, organizations must be cognizant of how important data is for verification, understanding, and quantifying decisions (Stobierski, 2019). According to a survey of over 1000 senior executives conducted by PricewaterhouseCoopers (PWC), businesses are significantly changing operations with data analytics (PricewaterhouseCoopers, 2018).

1.2 Statement of Problem

In recent years, data security laws have been enacted requiring organizations to take greater responsibility to maintain “reasonable” security controls of personal information from unauthorized access, destruction, use, modification, and disclosure (NCSL, 2019). , The current primary deterrents to prevent organizations from lax data security are compliance laws and regulations, which may lead to legal actions, fines, and imposed public notifications (Stanganelli, 2019). However, these are only a surface remedy for a far deeper problem to compel

organizations to take greater responsibility for stored data security because many critical elements are often overlooked. While several regulations focus on individual data security, organizations also need to consider all data (Brooks, 2019).

1.3 Research Questions and Hypotheses

To address the problem, several areas needed to be researched. At the core, three research questions related to data security were essential for this study:

1. If organizations comply with data privacy and security regulations, are they entirely securing stored data?
2. Does the social stigma of cyber incidents compel organizations to secure data?
3. Do data security laws, fines, and penalties compel organizations to implement stricter security controls for stored data?

To address these questions, the hypotheses were:

- H₀1: Compliance regulations do not miss any of the seven (7) principles of stored data security.
- H₁1: Compliance regulations miss at least one of the seven (7) principles of stored data security.
- H₀2: Social stigma of data breaches is not more critical to organizations than stored data security.
- H₁2: Social stigma of data breaches is more important to organizations than stored data security.
- H₀3: Avoiding fines and penalties is not more important than data security to organizations.
- H₁3: Avoiding fines and penalties is more important than data security to organizations.

H_x1 allows for a more objective or factual data set, whereas H_x2 and H_x3 produce a more subjective or opinion-based response. In both hypotheses, the datasets came from a random sampling of targeted individuals to determine if each of the hypotheses could be validated.

1.4 Objective and Contribution to the Discipline

This research examined current data privacy regulations and determined the strengths and weaknesses of stored data necessary for organizations and businesses to operate. To determine the mitigation techniques, it is vital to understand a data center's risks and gaps. This, in turn, may reduce data incidences and perhaps lessen the consequences by:

- Helping organizations simplify data and storage security
- Create a clear understanding of data privacy regulations
- Identify the areas that are lacking in compliance regulations as it relates to stored data

1.5 Limitation

The research was limited to individuals who were willing to contribute to this study and were willing to respond to a request to participate in a survey voluntarily. Individuals may be unlikely to disclose information that they feel is sensitive, embarrassing, or held to some form of Non-Disclosure Agreement (NDA). Also, individuals may not understand organizational security requirements or procedures due to outdated policies or lack of training. Lastly, while there are many laws and regulations related to data security, this research narrowed the scope to a subset of the most well-known and identifiable by working individuals in the United States. The limitations were summarized in chapter 5.

1.6 Dissertation Structure

This dissertation was divided into five chapters. It was organized following a numbered hierarchy structure of headings and subheadings for readability. Each chapter is outlined as follows:

Chapter 1: Introduction	<ul style="list-style-type: none"> → Purpose and goals of research → Background → Research questions and hypotheses → Limitations → Chapter orientation
Chapter 2: Literature Review	<ul style="list-style-type: none"> → Storage and data definitions → Article Analysis → The seven principles of stored data security → Security Controls for each stored data security → Regulatory Data Protection Models → Consequences and Penalties
Chapter 3: Research Methodology	<ul style="list-style-type: none"> → Population and sample size requirements → Scope and approach → Research tool development and IRB approval → Data collection and validation → Statistical and Data process analysis
Chapter 4: Results and Analysis	<ul style="list-style-type: none"> → Summary statistics → Stored data Security principles scoring results → Research question 1 analysis
Chapter 5: Conclusion	<ul style="list-style-type: none"> → Brief overview → Limitations → Findings and discussion → Summary and Future research
References and Appendices	<ul style="list-style-type: none"> → Work cited → Outline of 11 regulations utilized → Copy of developed survey → IRB approval → Specific category responses → Survey questions classification and measurement rubric → Glossary of statistical terms → Statistical reports

CHAPTER 2: LITERATURE REVIEW

At the core of most modern organizations is the data center and, for the most part, the technology that is relied on to maintain operations. Information technology is the gateway to the most precious commodity an organization owns: data (Madison, 2020). Data and storage cannot be siloed from one another as data security depends on its location, form, and how it is accessed.

2.1 Prevailing Definitions

Over the years, the definition of the term "datacenter" has varied slightly. However, for this paper, a data center is a physical or virtual space where information systems and data reside for an organization (Cisco, n.d.). The data centers can be a local or remote physical shared space or a managed service. The level of responsibility to support and maintain a data center can vary widely from organization to organization (Techopedia, 2017). With so many concerns for critical backend services, it is logical to consolidate the security, environment, power, and other requirements to a central location. Regardless of location or physicality, the datacenter is where applications and data are accessed (Techopedia, 2017).

Data has many forms, but for this research, data was classified as:

- Traditional structured or unstructured data
 - Structured data is stored in a predefined format and organized to be referenced, such as a database (Beal, n.d.)
 - Unstructured data is data in the form of flat files that are not in a predefined format, such as a text file, PDF, image, or typically another file that would be accessed by users directly (Komprise, 2009)
- Created, gathered, or the result of an organization's information systems, applications, or users and information stored in the form of either structured or unstructured data

- Virtual Machine (VM) data would include files for the metadata that define VM resources, virtual disks stored as a file that holds the OS, installed applications, and can have structured and unstructured data

For this research, the data types that were asked about were:

- **Personally Identifiable Information (PII)** – i.e., Individuals driver’s license, government id number, address, etc.
- **Protected Health Information (PHI)** – i.e., patients medical/health records, medications, treatments, etc.
- **Personal data** - i.e., age, gender, likes/dislikes, sexual orientation, religion, family, online social platforms info, diet, political views, pets, etc.
- **Employee information or data** – i.e., employees’ records, job roles, work schedule, vacation earned, salaries, bonus, etc.
- **Customer information or data** – i.e., customer information, order history, payment history, etc.
- **Financial data** – i.e., credit card data, investments, bank accounts, etc.
- **Student information** – i.e., grades, special needs, schedules, status, disciplinary actions, etc.
- **Data for minors** - i.e., children under the age of 18 years
- **Intellectual property** - i.e., trade secrets, procedures, designs, developed code, etc.

Data must be stored and secured somewhere and remain available for use. Data storage is a basic and fundamental function for a computer for fast access to information and resources (Khillar, 2018). The emphasis on data storage security for backend operations is identified as utility-based primary storage that is directly used for computer services. This type of storage can be in any of these forms:

- Storage Area Networks (SAN)

- Network Attached Storage (NAS)
- Direct Attached Storage (DAS)
- Software-Defined Storage (SDS)
- Object-Oriented Storage (OOS) or Object-based Storage Devices (OSD)
- Content Addressable Storage (CAS)
- Cloud Storage (examples such as OneDrive, Dropbox, Google Drive, Box)

Virtualization has paved the way for the next wave of datacenters. In the age of Cloud's "As-A-Service," datacenter storage has emerged into three primary classifications that embody functional requirements (Weins, 2018). The primary functional types were:

1. **Block storage** – storage presented to systems as raw-disk can be controlled and formatted by the host (Poojary, 2019)
2. **File storage** – hierarchical storage where files are organized under directories and are presented to the host as SMB/CIFS or NFS protocol (IBM, 2019b)
3. **Object storage** – storage that is separated into three parts: (Porter, Piscopo, & Marke, 2014)
 - a. The data (can be almost anything)
 - b. Expandable metadata – who/when created and any other relevant information
 - c. Global unique identifier

Archive storage has been identified as a type of storage. However, this is typically some type of near line cold storage used for a specific use-case, such as stagnant offline or long-term storage, and can be a subset of one or more of the three primary listed above (ISO, 2015). Also, data transport storage is related to the inter-communication of storage or data (Sarkar & Chatterjee, 2014).

For the most part, the three classifications represent most of the use cases needed for enterprise datacenter requirements. The simplification defines the nomenclature based on

functionality without the concern for the underlying technology or hardware. Regardless of the storage technology or active type, there is a need to standardize data and storage security.

2.2 Article Analysis

A systematic literature review of storage and data security was performed from September 2019 to December 2019, and over 100 articles were reviewed with 81 documents selected (see references for a detailed list). Documents and articles were found using the Dakota State University's Karl Mundt Library's research databases, which gave access to the following resources that were used throughout this literature review:

- ACM Digital Library
- IEEE Xplore Digital Library
- American Council for an Energy-Efficient Economy, or ACEEE
- Google Scholar
- InfoSecurityNetBase
- ProQuest Research Library
- National Technical Information Service (NTIS)

Well-known standards organizations were also researched. Notable and relevant sources included:

- Storage Networking Industry Association (SNIA) – Publicly available
- National Institute of Standards and Technologies (NIST) – Publicly available
- International Organization for Standardization (ISO) – Available through ANSI University Outreach Program
- Payment Card Industry Security Standards (PCI SSC) – Publicly available
- U.S. Department of Health and Human Services (HHS) – Publicly available

Articles were eliminated for review if they showed bias, were used as an advertisement for self-promotion or were written before 2005. Academic sources needed to be published and available on one of the library's research databases, and industry sources needed to be published by an industry professional organization.

Table 1 shows the top source publishers for articles on data or storage security:

Table 1 Top Article Sources

Top Sources for Articles and Papers						
<i>SNIA</i>	<i>ACM</i>	<i>IEEE</i>	<i>EBSCOhost</i>	<i>PCI SSC</i>	<i>InfoSecurityNetBase</i>	<i>Other</i>
29	12	8	7	5	4	16

By exploring only well-known organizations, the source material was scrutinized by experts within the industry. Academic sources were ensured to be peer-reviewed and published in recognized journals. Also, the publications can be seen with the number of articles selected by year (see Table 2):

Table 2 Years Published Grouped in three-year increments

Years Published				
<i>2005 - 2007</i>	<i>2008 - 2010</i>	<i>2011 - 2013</i>	<i>2014 - 2016</i>	<i>2017 - 2019</i>
3	15	5	22	36

The following notable results are observed:

- 35 of the articles were published through academic sources (43%)
- 43 articles were from industry sources (57%)
- Over 60% of the documents were produced after 2015

- The leading source (>35%) of publications was from SNIA

The publications' analysis was performed using a thematic approach, a method of analyzing qualitative data. It is usually applied to a set of texts where the researcher closely examines the data to identify common themes, topics, ideas, and patterns of meaning that come up repeatedly (Caulfield, 2019). When qualitative data, in this instance, publications related to data and storage security, were studied deductively, the themes or critical concepts were examined to show how this relates to the CIA Triad (Braun, Clarke, Hayfield, & Terry, 2019). As a result, the research's essential terms and ideas were tracked, and seven fundamental security principles emerged. Below are the security principles that are fundamental to data and storage security:

1. **Authentication:** Access control for validating that access to data and storage is allowed
2. **Authorization:** Access control for management and governance of authentication
3. **Privacy:** Ensure data and storage is isolated, encrypted, and allowed or decrypted by a valid source of authority
4. **Reliability:** Ensure data is accurate, and storage is durable and working as designed
5. **Verification:** Auditing, inspection, and analysis of data and storage
6. **Recoverability:** Ability for data and storage to return to a good or known state
7. **Accessibility:** Data and storage can be reached and usable as intended

These seven security principles can be correlated to the CIA Triad (See Table 3).

Table 3 Principles cataloged under CIA Triad

Confidentiality	Integrity	Availability
Authentication	Reliability	Recoverability
Authorization	Verification	Accessibility
Privacy		

2.3 Source Classification

Using this as a baseline for cataloging articles, a search of both academia and industry was conducted from various sources for data and storage security standards, best practices, and recommendations. The method was developed over three months and was tracked using Excel, where the articles were sorted by the seven defined security principles to count the breakdown. This became a working spreadsheet that produced a table that traced the driving themes of the research reviewed. The significant themes, keywords, topics, and ideas were noted during the process and were classified as the seven security principles that emerged. While many sources fell into multiple categories, the security principle was the single Classification or primary theme of the tallied article.

2.4 Article Mapping to CIA Triad

The findings were organized based on the document's primary category. This security taxonomy allowed a summation of all the reviewed documents, which formed an interpretation that produced observable patterns. The CIA breakdown is shown in Table 4, Table 5, and Table 6 below.

Table 4 Confidentiality

Confidentiality			
Totals:	30		
Security Principle	Authentication	Authorization	Privacy
Storage Security	4	6	4
Data Security	1	4	11
Academic	3	3	5
Industry	2	7	10

(Author, year)	1. (SNIA, 2018b)	1. (Butler, McLaughlin, & McDaniel, 2008)	1. (PCI SSC, 2015)
	2. (Schopmeyer, 2017)	2. (Tang et al., 2018)	2. (PCI SSC, 2017)
	3. (Hubbert, 2011)	3. (Hibbard, 2016)	3. (Sarkar & Chatterjee, 2014)
	4. (Daniel & Vasanthi, 2019)	4. (Willett, 2012)	4. (Krahn et al., 2018)
	5. (Park, Lim, & Kim, 2015)	5. (SNIA, 2015a)	5. (Hibbard, 2014)
		6. (SNIA, 2016b)	6. (Hibbard & Rivera, 2014)
		7. (SNIA, 2015c)	7. (SNIA, 2014)
		8. (McKay, Polk, & Chokhani, 2014)	8. (SNIA, 2018a)
		9. (ENISA, 2019)	9. (PCI SSC, 2018a)
		10. (Zhou, Varadharajan, & Gopinath, 2016)	10. (PCI SSC, 2010)
			11. (Schaffer, 2019)
			12. (Brandão, Davidson, Mouha, & Vassilev, 2019)
			13. (Zyskind, Nathan, & Pentland, 2015)
			14. (Wang, Yang, Duan, Guo, & Zhang, 2019)
			15. (Meslhy, Abd, Elkader, & Eletriby, 2013)

Table 5 Integrity

Integrity		
Totals:	26	
Security Principle	Reliability	Verification
Storage Security	12	6
Data Security	4	4
Academic	6	7
Industry	10	3
(Author, year)	1. (Jovanovic & Mirzoev, 2010)	1. (Hasan & Yurcik, 2006)
	2. (Butler, McLaughlin, & McDaniel, 2007)	2. (Vasilopoulos, Elkhiyaoui, Molva, & Onen, 2018)
	3. (Paik, Choi, Jin, Wang, & Cho, 2018)	3. (Zhu et al., Dec2010)
	4. (Hibbard, 2015)	4. (PCI SSC, 2018b)
	5. (Hibbard, 2011)	5. (Subha & Jayashri, 2017)
	6. (ISO, 2015)	6. (Hou, Yu, & Hao, 2019)
	7. (SNIA, 2017a)	7. (Schulz, 2011)
	8. (SNIA, 2016a)	8. (Kwon & Johnson, 2018)
	9. (SNIA, 2015b)	9. (Dell EMC, 2018)
	10. (SNIA, 2012)	10. (HDS, 2019)
	11. (SNIA, 2009)	
	12. (SNIA, 2010b)	
	13. (Gordan, 2019)	
	14. (Talib, Atan, Murad, & Abdullah, 2010)	
	15. (Dharma, Venugopal, Sake, & Dinh, 2013)	
	16. (IBM, 2019a)	

Table 6 Availability

Availability	
Totals:	24

Security Principle	Recoverability	Accessibility
Storage Security	7	5
Data Security	7	5
Academic	5	6
Industry	9	4
(Author, year)	<ol style="list-style-type: none"> 1. (Li, Qian, Chen, Hasan, & Shao, 2016) 2. (SNIA, 2019) 3. (SNIA, 2016c) 4. (SNIA, 2010a) 5. (McMinn, 2009c) 6. (McMinn, 2009a) 7. (McMinn, 2009b) 8. (Dutch, 2010) 9. (SNIA, 2017b) 10. (Schopmeyer & Somasundaram, 2009) 11. (Chang & Hao, 2009) 12. (Jian-hua & Nan, 2011) 13. (Wang & Cheng, 2018) 14. (Bollinger, Enright, & Valite, 2015) 	<ol style="list-style-type: none"> 1. (Zhou, 2014) 2. (Chen & Zadok, 2019) 3. (Carlson & Espy, 2017) 4. (SNIA, 2008) 5. (Fuxi & Yang, 2015) 6. (Rouse, 2019) 7. (BlockApps, Dec2017) 8. (Xu, 2018) 9. (Zheng, Li, Chen, & Dong, 2018) 10. (Veleva, 2019)

Only the author(s) and year published were recorded in the three tables; however, all source documents were cited in the reference section.

2.5 Results of Classification

It should be noted that the Classification for each article and document was performed autonomously to avoid biasing the results. The breakdown based on the CIA Triad was relatively evenly distributed with 30% Availability, 33% Integrity, and 37% Confidentiality. It is also notable that neither the academic nor industry articles focused on one specific security area and were split evenly across data and storage security (see Table 7).

Table 7 Source Type by Concentration of Security Interest

Source Type	Storage Security	Data Security
Industry	54.35%	45.65%
Academic	57.14%	42.86%

Understanding if an article was published by an academic or industry source, an inference could be suggested for which area was of most concern. Using the frequency summary from the above tables, a graphic was created to cross-reference the sources' classification and each of the seven principles. The graph below compares academic and industry articles cataloged by the seven defined security principles (see Figure 4).

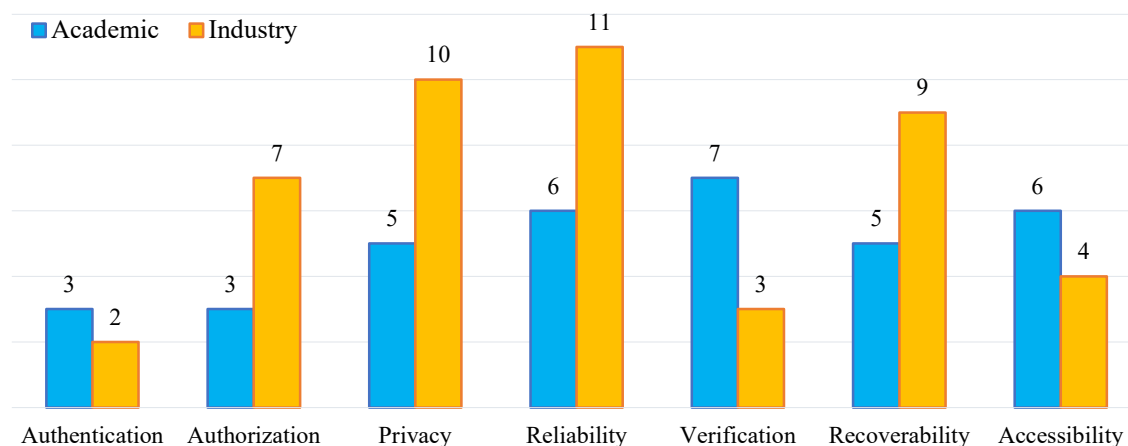


Figure 4 Security Principles Shown by Academic and Industry Articles

The above graph illustrates that the industry articles prioritized four security principles: reliability (11), privacy (10), recoverability (9), and authorization (7). Whereas academic sources were more leveled focus on verification (7), reliability (6), accessibility (6), privacy (5), and recovery (5). The gaps from the industry were accessibility (4), verification (3), and authentication (3). While for academics, the gaps were only for authorization (3) and authentication (3).

In addition to comparing academic to industry, the articles were categorized as related to data or storage security. The comparison showed a significant majority of articles related to data favored privacy (11), where storage favored reliability (13). Figure 55 shows a comparison based on storage and data security.

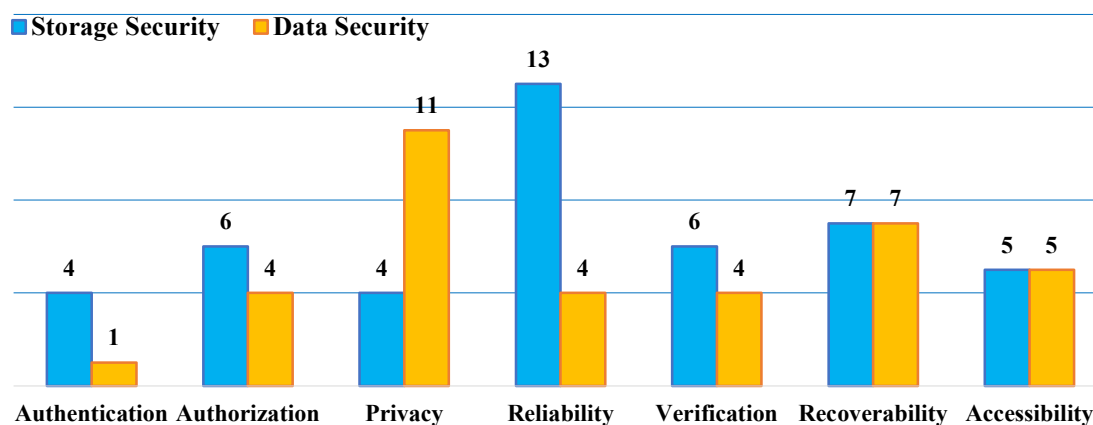


Figure 5 Security Principles Compared by Data and Storage Articles

Based on the observations of the articles analyzed, the findings indicated that gaps existed related to authentication as a primary theme to data and storage security. What was clear was that data security's primary theme was related to privacy and storage security focused on reliability. When closely examining the two graphs, the connection can be seen, and the industry primarily focused on privacy for data security and reliability for storage security.

2.6 Principles of stored data security

Every organization's information security, access controls, and safeguards are subject to confidentiality, integrity, and availability (Agarwal & Agarwal, 2011). This security policy development model is otherwise referred to as the CIA Triad. This model is essential in modeling security around data storage and ensuring an organization complies with data privacy and protection measures. Data privacy and security builds trust between organizations and their customers, and the CIA triad is a non-industry specific tool towards this goal (Imam, 2019).

2.6.1 Confidentiality

Confidentiality is the aspect of ensuring privacy for data that is stored, in transit, or over a network. This prevents unauthorized access to sensitive data while simultaneously making it available to intended users (Agarwal & Agarwal, 2011). Through access control and privacy, confidentiality can be attained. Confidentiality involves allowing data access to authorized users and restricting unauthorized persons. It contains three components: authorization, authentication, and privacy (Agarwal & Agarwal, 2011).

2.6.1.1 Authentication

Authentication refers to any process which verifies that someone is whom they say they are. This mainly involves a username and password and includes enhanced technologies such as biometrics (i.e., retina scan, fingerprinting, or voice recognition) (Fruhlinger, 2020).

Username/password combination remains the most prevailing means of authentication (Siponen, Puhakainen, & Vance, 2019). Passwords are strengthened by making them more complicated, updating them regularly, disallowing the use of previous passwords and other techniques, making them more difficult to crack (NortonLifeLock, n.d.).

Authentication can be further enhanced to improve security using strong multi-factor authentication. A good example is two-step verification: after inputting the correct credentials, such as username and password, a predetermined or trusted destination such as a device via text, email, or voice message sent with a code, link, or other means confirm validity (Garun, 2019). These and other methods are intended to establish the identity of the individual requesting access to data. Further, encryption is implemented to enhance security during authentication. In the

process, the password and username are scrambled in an unintelligible text (ciphertext) that only the receiver can understand and decode (Krzyzanowski, 2009).

2.6.1.2 Authorization

Authorization determines users' permissions or privileges in a system (Fruhlinger, 2020). Permissions are defined by creating, viewing, altering, or deleting data that requires security controls. Authorization controls are vital in protecting data against unlawful, unauthorized, or accidental incidents (ISO, 2015). The system uses a defined access policy to grant or deny requests made by authenticated parties (Nelson, 2017). For instance, in a multi-user system that combines different departments in an organization, authorization differentiates the data that can be accessed by human resources and accounting.

Access controls are the baseline protection of data against unauthorized access, modification, and time out after a period of inactivity within a given authenticated active user session (Carnegie Mellon University, 2020). Authorization also clarifies the data that department heads can access and alter the subordinate's rights and privileges; for example, read-only and read-write access (Hoven, Blaauw, Pieters, & Warnier, 2019). Lastly, authorization should encompass access revocation to disable access when users no longer need it or leave a program (Temple, 2016).

2.6.1.3 Privacy

The privacy of stored data can be achieved by enforcing pertinent policies, procedures, and defense mechanisms (Mulligan, Freeman, & Linebaugh, 2019). Through risk evaluation and management, access restrictions, and data encryption, unauthorized access to private or susceptible data can be protected multilaterally. When required to share confidential

information, such as contracts or business agreements, a Non-Disclosure Agreement (NDA) is used when authorizing sensitive information is necessary while creating a legal obligation to privacy (Alexandra Twin, 2020). Privacy is considered a fundamental right, and in the digital age, one of the most challenging aspects of security to control. For this reason, many compliance regulations include requirements for consent and allow for the revocation of usage or storage of personal data upon request (GDPR.EU, 2018b).

Encryption is one of the most efficient technical controls for ensuring stored data privacy (Baig, 2020). At its most basic, it is the encoding of data into a cipher such that a private digital key is needed to decode back to readable form (E. Hibbard, 2016). If implemented at the physical level of the storage media or the file system's logical level, encryption of the data is independent of authenticated access (Kumar, Rawat, Jasra, & Jain, 2009). Relying too much on one form of encryption solely can lull organizations and individuals into a false sense of security (Vandersreen, 2019). This is not to say that encryption is unnecessary; on the contrary, modern cryptographic techniques are fundamental for ensuring data security (S. Butler, 2018). Nonetheless, cryptography techniques need updating as technology shifts since they may pose risks as they become outdated by faster computers and newer threats (S. Butler, 2018).

2.6.2 Integrity

Stored data should remain original, accurate, and unaltered either by mistake or maliciously (Fruhlinger, 2020). Data changes must be tracked somehow in intentional or unintentional unwanted actions to data or the underlying storage. Ideally, data should not be altered by an authorized party and should remain in its original state; however, incidents occur. For this reason, integrity comprises two parts: reliability and verification.

2.6.2.1 Reliability

The reliability of data can be evaluated depending on its accuracy, consistency, and durability. Corruption and degradation prevent retrieval or recoverability of data in its initial form (Blum & Singh, 2017). It is inevitable, regardless of whether it is legacy media or modern flash, that storage can fail. Storage must ensure that no single-points-of-failure (SPOF) and misconfigurations exist at the file or disk level (ISO, 2015).

Technologies such as Redundant Array of Independent Disks (RAID) allow for hardware failure mitigation. RAID is a storage virtualization strategy that distributes data across multiple disks through disk striping and disk mirroring (Rouse, Sullivan, Posey, Diamantis, & Yamamura, 2020). While RAID and other storage protection technologies may have many different levels to protect against one-to-many failures, it is essential to understand that the primary goal is to maintain data reliability (SNIA, 2016b). Storage durability is achieved by implementing data redundancy to protect against degradation or corruption (W. Li, Yang, & Yuan, 2015).

Data reliability is an essential aspect of data security as organizations and individuals trust the integrity of information stored. With the continual growth of data, reliability, accuracy, and completeness require error checking and validation (Naeem, 2020). As a measure of administrative control, reliability must enable a means to correct discrepancies when discovered (Harkness & Black, 2020).

2.6.2.2 Verification

Verification in data storage security involves monitoring, logging, and recording stored data viewed, added, modified, or deleted by system processes and user access (Kent & Souppaya, 2006). Data integrity is a challenging and arduous job in the age of cloud-based

technology and globalization. Users and organizations need to rely on stored data that may not be local or outside of physical control (C. Wang, Wang, Ren, & Lou, 2010).

Auditing is vital as it enables accountability of actions carried out on the data and detected problems with access controls enforcement (Bowen, Hash, & Wilson, 2006). Thorough auditing is performed periodically to ensure all security policies and controls are tested up to data security in the event they are needed. Keeping records on audit logs, authorization, and authentication events for verification are crucial in proving compliance with regulatory authorities, users, and other stakeholders (Bowen et al., 2006).

Verification of data security measures in terms of efficacy, completeness, and accuracy is vital in inspiring confidence users, storage services providers, and regulatory authorities (Neuhäuser, Lehmann, Nonnemacher, & Stausberg, 2006). This activity's basis lies in data classification according to its level of sensitivity, value, and importance, as outlined by data security policies (Carnegie Mellon University, 2018).

Data classification also encompasses understanding the threat and impact of both the organizations and users in data breaches, losses, or unauthorized alterations. The primary data classes include restricted, private and public data (Carnegie Mellon University, 2018). Governments also classify their data as either top-secret, secret, confidential, or unclassified ("Classified Information," 2020). It is important to note that auditing and protection are intensified depending on the levels of risks associated with the data. Classification levels that pose the highest risks, for instance, restricted data or top-secret data, get more attention.

Lastly, over recent years, the governing body, public and individual notifications, or public data incidents and breaches have become a vital deterrent for validation (Schneider,

2009). Neglect of the responsibility to notify regulators, stakeholders, and affected customers of data breaches can damage an organization. Notifying governing authorities can help mitigate and prevent future prevention measures and identify individuals who have been affected (Burnette, 2018).

2.6.3 Availability

Availability in data storage security means that authorized users can access data resources when they need them (Imam, 2019). Different factors that threaten data availability include hardware and power failures, natural disasters, malicious attacks, and human errors (Walkowski, 2019). Data storage that satisfies the CIA Triad's availability requirement can be gauged on two principles: recoverability and accessibility.

2.6.3.1 Recoverability

Recoverability describes a principle in which data storage is secured in times of data losses so that organizations can retrieve and maintain data following a disaster (Rouse et al., 2020). To achieve this principle, organizations require adequate data backup and recovery systems to return normal operations after disasters.

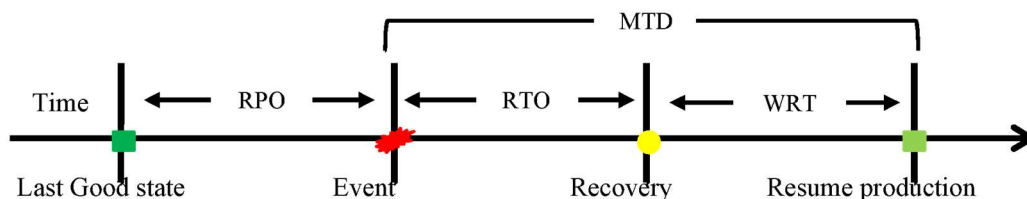


Figure 6 RPO, RTO, MTD, and WRT (Marek.Z, 2013)

When considering data storage recoverability, organizations should consider four key parameters: Recovery Point Objective (RPO), Recovery Time Objective (RTO), Work Recovery Time (WRT), and Maximum Tolerable Downtime (MTD) (Marek.Z, 2013).

RTO is a metric that evaluates the amount of time required to fully recover data in the case of data losses and resume business operations (Dennis G, 2020). This time is an indicator of how long an organization can survive following a disaster before normal activities are restored. On the other hand, RPO measures the maximum gap between the last backup and a disaster experienced that still means a business can sustain continuity (Dennis G, 2020). In other words, RPO determines how often organizations should perform data backups. The backup aims to create a copy of data that can be retrieved in case of a core data failure event. These parameters are vital in drafting business continuity plans since their deterioration could impede business operations' normal flow.

Other advanced data protection systems include replication and redundant data storage. Redundancy enhances fault tolerance since secondary data storage remains available if the primary storage fails (SNIA, 2017b). Due to the high dependence on business-critical IT services, their unavailability could cause imminent financial loss and negatively impact the social construct (Carnegie Mellon University, 2020). Regarding the safety of data, enhancing data recoverability in case of losses is vital. Therefore, it is crucial to implement backup and recovery and redundant systems that are efficient enough to ensure the resumption of operations after a disaster (ISO, 2015).

2.6.3.2 *Accessibility*

Accessibility is a component of availability that dictates that stored data should be reachable and usable as required at any time (Tozzi, 2020). Multipath I/O is a technique that specifies multiple physical paths between a system and data storage devices to enhance fault-insusceptibility and performance (Rouse, Cook, & Wigmore, 2012). This may be achieved through data buses, bridges, controllers, or switches connecting the system and data storage. This technique enhances redundancy and load balancing for storage devices and is very useful in a host bus adapter (HBA) malfunction (Rouse et al., 2012). Load balancing refers to efficiently distributing incoming traffic and requests across multiple data storage devices (NGINX, 2018).

Another technique of enhancing accessibility is the network or network interface controller (NIC) teaming. This process involves linking multiple network cards to enhance redundancy, performance, and load balancing (Collins, 2020). In the case of faults, the target hosts detect fault conditions and automatically reroute traffic through a different NIC in the pathway (Collins, 2020). This ensures that data and storage are accessible to authorized users when required. Through NIC teaming, the system allows for interface bonding of multiple physical network adapters to increase performance by link aggregation (routing traffic over multiple network adapters) and enhance fault tolerance (Agrawal, 2016). One network adapter's failure can be mitigated as the system dynamically reconfigures functional NICs in the bond.

High availability (HA) clustering is another way of enhancing accessibility. This involves a group of computers that run high availability software to enhance redundancy and continued service if one malfunctions. This ensures that services remain available following faults and device failures (Heder, 2014).

Finally, the most common threat to accessibility out of malicious attacks is the denial-of-service (DoS) or distributed denial of service (DDoS) attack. These types of attacks prevent authorized users from accessing information systems and data, which costs time and money for the company (CISA, 2019, p. 201). There are several types of DDoS attacks, but tools and techniques can help defense (DNSStuff, 2019). Measures include network/threat monitoring tools, firewalls, Web Application Firewalls (WAF), anti-virus/malware, and other network protection controls (Gupta, Perez, Agrawal, & Gupta, 2019).

2.7 Security Controls

Modern storage and networks have standard security solutions with many common characteristics, capabilities, and features. This, in turn, has enabled technical advances for organizations of all sizes, budgets, expertise, or locations. The seven principles provide a clear delineation of the areas for data protection and security. However, with so many security techniques, it can be challenging to determine if current implemented technologies, procedures, and restrictions cover all seven.

A clear view can be formed by listing the seven principles and cross-referencing them to physical, administrative, and technical controls to determine areas that may be underrepresented. Table 8 shows how this can be performed with ~120 security controls. However, evolving technologies and techniques make this only a snapshot in time. Nevertheless, in this method, by cataloging the defense measures, an unbiased view can be formed on which area of data security needs attention.

Table 8 Seven Principles cross-referenced to Security Control classifications

Principle	Security Controls		
	Physical	Administrative	Technical
Authentication (14 Controls)	<ul style="list-style-type: none"> Physical security verification measures Biometrics Token-based device User registration and de-registration 	<ul style="list-style-type: none"> Username/password Required to update passwords regularly Complex passwords Passwords cannot be reused Assign unique IDs to all users Admin passwords are changed regularly 	<ul style="list-style-type: none"> Two-factor or multi-factor authentication CAPTCHAs (validate human access) Single sign-on (SSO) User and Entity Behavior Analytics (UEBA)
Authorization (15 Controls)	<ul style="list-style-type: none"> Self-service web administration Dedicated helpdesk support Restrict physical access to classified data Restrict the use of external storage (USB, cloud storage, etc.) User access provisioning 	<ul style="list-style-type: none"> Management of privileged access rights User role management Automation of providing access (request) User rights, access, and password management Management of authentication information of users 	<ul style="list-style-type: none"> Web Application Access Control Access Management Virtual Directory Active Directory, LDAP, NIS+ Identity and Access Management (IAM) Password management system
Privacy (20 Controls)	<ul style="list-style-type: none"> Encrypted drives and storage Secure user areas Restrict the use of video recording devices Restrict the use of personal email and IM 	<ul style="list-style-type: none"> Opt-in or opt-out Cyber Threat Intelligence Federal Information Processing Standards Key management Security patch management Asset Management Restrict physical and remote system assets Restrict the transition of data to untrusted destinations Clear desk and clear screen policy 	<ul style="list-style-type: none"> Encryption technologies Public Key Infrastructure (PKI) File-level permissions Anti-Virus Anti-ransomware Security patching automation privacy information management system (PIMS)
Reliability (14 Controls)	<ul style="list-style-type: none"> Security and privacy leadership Well documented designs for security and privacy Limit access to information by time or need 	<ul style="list-style-type: none"> Privacy and strategic security planning Security and privacy policies and procedures Ability update and modify inaccurate information Operational procedures and responsibilities Change management 	<ul style="list-style-type: none"> Standard security and privacy tools Content monitoring/filtering tools Event Log Analysis Tools Security Information Management (SIM) RAID Erasure Coding
Verification (15 Controls)	<ul style="list-style-type: none"> Suppliers and IT risk Management Management of removable media Media handling and disposal Regularly test security 	<ul style="list-style-type: none"> Classification of Information Standard terms for security and privacy Fraud Investigation Compliance with IT regulatory requirements Risk, readiness, impact, and 	<ul style="list-style-type: none"> Centralized logging and monitoring Security Event Management Endpoint Detection and Response (EDR)

	<ul style="list-style-type: none"> • systems • Inventory of assets 	<ul style="list-style-type: none"> • other analysis or audits • Logging and monitoring • Vulnerability management 	
Recoverability (19 Controls)	<ul style="list-style-type: none"> • Recovery data center • Multiple zones • Physical media transfer • Protecting against external and environmental threats • Dedicated storage only networks (SAN, NAS, etc.) • Management of removable media • DR team 	<ul style="list-style-type: none"> • Ability to restore data in the event of loss or damage • Disaster Recovery Plan (DRP) • Business Continuity Plan (BCP) • Notifications and call chain • Backup and recovery policies and procedures • Information security continuity • RPO, RTO, MTD, and WRT 	<ul style="list-style-type: none"> • Remote and local data replication • File, application, or block-based replication • Synchronous / Asynchronous replication • Disk or tape backup system • Continuous Data Protection (CDP) solutions
Accessibility (22 Controls)	<ul style="list-style-type: none"> • Network redundancies • System network isolation • Use badges or other identity validation • Human security • Secure data center • Physical security perimeter • Anti-tailgating • IR Team 	<ul style="list-style-type: none"> • Access control policies • Access control models • User access management • Review of user access rights • Removal or adjustment of access rights • Secure log-on procedures • Network security management • Network Access Control (NAC) 	<ul style="list-style-type: none"> • Intrusion Prevention Systems (IPS) • Intrusion Detection and Prevention Systems (IDS) • Transaction location checking (i.e., IP checks, device recognition) • Load balancers • Firewalls and WAF • Network Monitoring tools

2.8 Regulatory Data Protection Models

Data protection is essential for both individuals and organizations, transparency in how information is collected, how it is stored and protected, and how organizations adhere to security policies is a cornerstone to building trust among consumers and shareholders (Peters, 2020). The need for data security protection has resulted from such factors as demand by market trends, need for compliance with regulatory requirements, advances in data storage, privacy, and security trends, among others (Fuller, 2019). There is just no single source or approach for data protection (Peter P. Swire & DeBrae Kennedy-Mayo, 2018).

According to the Economist Intelligence Unit (EIU), a more significant number of consumers want more transparency and control in handling their private data (Fearn, 2018). This

forms the basis of the core concept of data privacy protection (Eliezerov, 2020). The continued improvement of awareness on the importance of their privacy has led to an emphasis on data security. Another contributor to the need for data security is policy compliance, either imposed legal requirements or self-regulation. Such legislation includes the Consumer Data Protection Act, the General Data Protection Regulation (GDPR) in Europe, and the Data Care Act (Achary, 2019). To ensure adherence to these legal requirements, data companies establish compliance departments to ensure that their organizations operate within the defined regulations' precincts or abide by their privacy policies.

Advances in technology call for shifts in stored data protection to ensure consumer data remains protected. The imminent capabilities of telecommunications and computing technologies and how they impact the data exchange and flow raise the need for personal privacy protection and awareness (OECD, 2011). As technology shifts to the cloud, enhance techniques for secured data protection standards are needed (Pottier & Menaud, 2017). This means that organizations may need to get creative. For example, instead of the traditional way of preserving the user's actual email address, Apple announced it would generate new email addresses on behalf of its users where emails will be sent to and then redirected to the actual email addresses (Apple, 2020). Such shifts in anonymity in technology can invoke better ways to handle consumer data in areas ranging from home deliveries, telephone communications and the transfer of funds (Eliezerov, 2020). For now, the best means for data protection is to start with a baseline for minimum protection standards.

2.8.1 Ensuring compliance

The goal of compliance regulations is to ensure the security and privacy of data. Despite strict enforcement, data breaches and incidents still occur. Far too often, regulations neglect one

or more areas for a complete examination of security controls for verification of detection, prevention, or correction of security incidents (Swire & Kennedy-Mayo, 2018). Understanding which regulations apply to an organization can be confusing because laws and regulations cover different data sets and their use (Thomson Reuters, 2020).

2.8.2 Comprehensive Data Protection Model

Data security enables us to make choices on who has access and how our information is used. Comprehensive data protection laws are consequentially crucial in protecting this right. These laws regulate how government and private organizations collect, use, and share personal information in all places employing this data protection model (Swire & Kennedy-Mayo, 2018). The comprehensive models look to protect a specific population of people instead of a specific data type (Swire & Kennedy-Mayo, 2018). The most well-known example is the European Union's GDPR, which applies to all European citizens (Fearn, 2018). GDPR outlines policy requirements and stiff penalties for institutions that fail to comply with the European Union's data protection policies (Saltis, 2020). Independent institutions, known as the Data Protection Authority (DPA), exist to enforce the outlined laws and policies (European Commission, n.d.). DPAs hold investigative and reformatory powers to ensure the application of data protection law. As of 2020, the United States does not have a comparable regulation; some states follow the EU's lead. The California Consumer Privacy Act (CCPA), signed into law on June 28, 2018, protects California consumers' personal information across industries and is considered the closest equivalent (Becerra, 2018).

GDPR and CCPA have positively impacted several existing problems (DataGuidance, 2018). The assumption being, improving data privacy measures for both public and private organizations would limit the potential for data loss, breaches, and other incidents (SIRE, 2019).

Having a common framework ensures that data protection is standardized; thereby, consistency in compliance enforcement is realized. Finally, comprehensive data protection models have refurbished customer's trust in organizations by assuring them that their personal information is secure from unwarranted access (McGavisk, n.d.).

2.8.3 Sectoral Data Protection Model

As opposed to the comprehensive data protection model where data protection is applied across all industries, some countries adopt a model that employs data protection legislation when specific industries and circumstances require it. The United States uses this model of industry-specific rules. For example, the United States passed the Health Insurance Portability and Accountability Act (HIPAA) in 1996 (Woodard, 2004). This act issued directives on standards to secure electronic health information by the Department of Health and Human Services (HHS), and health services providers had to comply. Consequently, HHS created HIPAA regulations, including the Privacy Rule of 2003 (OCR, 2015). The Privacy Rule provides security for any personally identifiable health information or Protected Health Information (PHI).

In the sectoral approach, many industries are regulated by overlapping data security laws, both federal and state, resulting in an inconsistent and complex compliance enforcement process (Solove, 2015). While healthcare data is recognized as protected by HIPAA, the regulation does not apply to all organizations (OCR, 2015). Health data will be regulated when held by a hospital, insurance company, or school but not by a tech company (OCR, 2015).

2.8.4 Co-Regulatory and Self-Regulatory Models

Co-regulation defines a data protection strategy where both government and industry cooperate to regulate data protection (Hirsch, 2013). The industry gets involved by developing

codes of stored data protection and privacy in line with the law's requirements that government can enforce. A good example is the US Children's Online Privacy Protection Rule (COPPA) signed into law in 1998. While Senators Bryan and McCain spearheaded the creation of a child safety and protection law, the legislative work process included industry, government, and civil society groups collectively developing and executing principles guiding children's online safety (A. Cohen, 2019). The COPPA Rule developed can then be enforced by the US Federal Trade Commission (FTC). Co-regulation inputs the elements of accountability and consumer trust while maintaining self-regulation (A. Cohen, 2019).

The self-regulatory framework offers several advantages. First, it is more economical than government regulation as organizations can independently grow accustomed to their own needs and efficiently abide by their internal policies (Domingo & Villar, 2018). Second, government regulations can hamper innovation, especially when they fail to adapt to the forever changing technological landscape. Arguably, by ensuring competition, companies can foster innovation to realize the most secure data protection technologies (Domingo & Villar, 2018).

An example of the self-regulatory model is the Network Advertising Initiative Self-Regulatory Code of Conduct. This code limits the types of data that member companies can use for advertising while enacting stringent restrictions on member companies' collection, use, and sharing of data used for tailored advertising (National Advertising Initiative, 2018). To ensure strict compliance mechanisms that include sanctions (National Advertising Initiative, 2018).

2.8.5 Selected Regulations

Regulatory compliance can vary based on industry, type of data, or population. As part of this research study, eleven regulations were selected for analysis. The criteria for selection

were based on regulations that were well-known within the United States. The following list was the regulations referenced for this study:

1. BASEL II
2. California Consumer Privacy Act (CCPA)
3. Family Educational Rights and Privacy Act (FERPA)
4. Federal Financial Institutions Examination Council (FFIEC)
5. Federal Information Security Management Act (FISMA)
6. Federal Risk and Authorization Management Program (FedRAMP)
7. General Data Protection Regulation (GDPR)
8. Gramm–Leach–Bliley Act (GLBA)
9. Health Insurance Portability and Accountability Act (HIPAA)
10. Payment Card Industry Data Security Standard (PCI DSS)
11. Sarbanes–Oxley Act (SOX)

The regulations were sorted alphabetically. [Appendix A](#) shows a summary outline of each of the eleven regulations organized by the following information collected:

- Regulation name and acronym
- Official website
- Year implemented or enacted
- Data protection model
- Industry or purpose
- Governing Body or enforcement authority
- Data Types protected
- Brief description of the regulation
- List of requirements
- Non-Compliance consequences or penalties
- Name data security controls
- Which of the seven stored data principles are covered

The objective was to summarize the security controls, requirements, and consequences to determine if the regulations mapped to the seven principles of stored data security. Table 9 showed the analysis of each of the eleven regulations outlined in [Appendix A](#) and used for the research study.

Table 9 Regulations mapped to the seven principles of stored data security

Regulation	Authentication	Authorization	Privacy	Reliability	Verification	Recoverability	Accessibility
FFIEC	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FISMA	Yes	Yes	Yes	Yes	Yes	Yes	Yes
FedRAMP	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CCPA	Yes	Yes	Yes	Yes	Yes	No	Yes
FERPA	Yes	Yes	Yes	Yes	Yes	Yes	No
GLBA	Yes	Yes	Yes	Yes	Yes	Yes	No
PCI DSS	No	Yes	Yes	Yes	Yes	Yes	Yes
SOX	No	Yes	No	Yes	Yes	Yes	Yes
BASEL II	No	No	Yes	Yes	Yes	Yes	No
HIPAA	No	No	Yes	Yes	Yes	Yes	No
GDPR	No	No	Yes	Yes	Yes	No	No

The above table was sorted based on the number of stored data security principles that were evident. It was noteworthy that the top 3 regulations were federally regulated, and the least complete was GDPR. Of the eleven regulations, four were missing one principle. All of the regulations had security controls or requirements for reliability and verification, and only SOX didn't have any rules or management related to privacy.

2.9 Consequences and Penalties

There are many consequences associated with data security breaches. This includes loss of finances and customers, civil lawsuits and litigations, and monetary penalties. Different regulations define baselines for data privacy and protection policies that should be enacted to ensure that systems have matching security measures to protect data. In the cases of information security breaches due to the failure of meeting these basic levels of protection, then organizations

in violation are subject to monetary penalties and other following consequences. Understanding these consequences is critical in emphasizing the importance of data privacy protection.

Depending on the nature of data and privacy breaches, data protection authorities may issue administrative fines or monetary penalties may be issued by courts during litigations. The penalties vary among different regulations and depend on the nature, severity, and duration of violations. For example, the Information Commissioner's Office (ICO) could impose fines as high as 25 million dollars or higher under the GDPR, depending on the violations. US Department of Health and Human Services (HHS) recommends fines of up to \$7500 per violation under HIPAA (OCR, 2015).

Data breaches harm customer relationships, which can lead to customer loss. Data indicates that customers' adverse perception in the face of a data security breach could trigger customer abandonment to competitors. In competitive industries with multiple organizations offering similar services, customers shift to the competitors to ensure their data safety (CyberInsureOne, 2019). Loss of customer trust and their investments consequently results in revenue losses by the affected organizations. The massive loss of fines in the aftermath of a breach comes in the forms of penalties and fines, cost of response and recovery, cost of investigation, and other associated losses that scare investors away. Also, businesses whose core income is generated online, such as e-commerce, are directly impacted by immediate revenue losses until normal operations can be restored (Ponemon Institute & IBM Security, 2020).

Apart from financial losses, data breaches have a long-term consequence of reputation damage . While many organizations' first response to data breaches is to conceal the information to minimize reputation harm, the strategy can exacerbate damage once the cover-up is exposed. A trust bonds customer relationship as any other partnership, and therefore, reputational damage

results in loss of existing and potential customers and investors, particularly in a competitive market space (Clearswift, 2019).

Organizations affected by data breaches could face civil lawsuits from affected customers, business partners, and government agencies. In most cases, organizations have to prove that the data breaches were not a result of negligence and that the best security measures and practices had been enacted (CyberInsureOne, 2019). In the Business Continuity Institute Horizon report 2018, data breaches were ranked second as the most disruptive to business operations for the past three years (BCI, 2018). Cybersecurity breaches' impacts include loss of revenue, and customers and lawsuits can threaten business continuity (Clearswift, 2019). Several studies have emphasized the damaging effect data breaches have on revenue and investments (Klebnikov, 2019). They are reporting stock price returns to deteriorate in the years following a data breach, followed by stockholder demands for enhanced security controls to prevent future incidents (Klebnikov, 2019).

According to a Ponemon Institute study performed in 2019 and 2020, the average cost of data breaches in the United States is more than \$8 million and can taking over 6 months to identify and isolate (see Table 10). The report also showed that the industry most affected was healthcare. In the light of these reports, the Institute recommends compliance with industry data privacy and protection policies to prevent incidents that pose significant risks to businesses' finances and continuity (Ponemon Institute & IBM Security, 2020).

Table 10 U.S. cost for data breaches 2019 – 2020 (Ponemon Institute & IBM Security, 2020)

Averages	2020 Results	2019 Results
The total cost of a breach	\$8.64M	\$8.19M
Time to identify and contain	237 days	234 days
Security automation deployed	76% of orgs	58% of orgs
Costliest industry	Healthcare	Healthcare

As more industries face scrutiny by administrative authorities, organizations are embarking on data protection against breaches to avoid punitive measures enacted upon them by authorities. However, most companies understand the threat posed by cybersecurity on their businesses, including long-term distrust by customers and halt of operations. By this understanding, organizations continue to establish policies that safeguard their businesses against data breaches (Verizon, 2020).

2.10 Summary

In this chapter, an in-depth review was conducted on storage and data security that produced a new framework to analyze security concerns and controls for stored data. Each of the seven principles was described in detail, and a table was developed that cross-referenced each principle to the three control layers: physical, administrative, and technical. Over 120 techniques, policies, and tool categories were identified and outlined for modeling and reference. After the model was completed, additional research was performed on the data protection models, which helped identify 11 regulations and laws that would be easiest for people to recognize. The last section reviewed the common imposed penalties and consequences because of not protecting personal and private data. The amalgamation of the research fostered the development of the primary research survey tool described in chapter 3.

CHAPTER 3: RESEARCH METHODOLOGY

This chapter describes the research methodology chosen to explore the hypotheses for this dissertation. The research aims to measure the strengths and weaknesses of various data

regulations and motivators to determine which factors compel organizations to implement more robust stored data security.

Due to this research's nature, a quantitative approach was applied for critical analysis of professional and industry opinions and understood day-to-day issues related to data security and different IT operations. The research questions being addressed require a mix of professional opinions and quantifiable data sets. The quantitative addresses a better understanding of the factors or variables that may influence the outcome related to data storage incidences (Creswell, 2014). The data for this research was derived from an anonymous online survey as described in detail below.

3.1 Population and Sample Size Requirements

Sample size determination involves making inferences about the population based on the sample. It solely consists of choosing the number of observations or replicates to include in a statistical model (Noordzij, Dekker, Zoccali, & Jager, 2011). In most designs, the different sample sizes are determined based on time, cost, convenience in collecting the data, and the sample's need to offer sufficient statistical power (Noordzij et al., 2011). Sample sizes in detailed studies such as stratified surveys have different sample sizes for each stratum (Hayes, 2020). By the very nature of conducting an anonymous survey, it should be expected to see an imbalance of sample sizes among different groups of participants due to simple randomization (Vanhove, 2015). Variable sample sizes may result in random requests for participation. However, this only adds to validation as it avoids sample bias (Kahan, Rehal, & Cro, 2015).

Cochran (1977) developed a formula used to calculate a representative sample for the population proportions (Isreal, 2005). In this research and with a broad inclusion criterion, such

as the “all working adults residing in the United States,” the estimation for the population size can be estimated at over 100 million (US Bureau of Labor Statistics, 2020). When the population is large, Cochran’s formula for calculating the sample size is defined as:

$$n = \frac{Z^2 pq}{e^2}$$

Whereas,

n = desired sample size

Z = selected critical value of desired confidence level

If $X \sim N(\mu, \sigma^2)$ then $Z = \frac{X - m}{s}$ is a standard normal variate i.e., $Z \sim N(0, 1)$

p = estimated proportion of an attribute that is present in the population = 50%

q = estimated proportion of an attribute that is not present in the population

$q = 1 - p = 1 - 0.5 = 0.5$

e = the desired level of precision/ degree of accuracy (margin of error) usually set at 0.05

Figure 7 Cochran’s formula for calculating sample size (Cochran, 1977)

When computing a confidence level for an estimate from a large population, it is statistically acceptable to use the Z -table for the desired confidence interval needed (See Table 11).

Table 11 Common Z -Table for standard Confidence Levels (LTCC, 2009)

Confidence Level	Z -score (\pm)
70%	1.04
75%	1.15
80%	1.28
85%	1.44
90%	1.65
95%	1.96
96%	2.05

98%	2.33
99%	2.58

The confidence level is the area under the standard curve. The Z-score represents this value for 95%. Table 11 shows the Z-score values for confidence levels 70% to 99%. Using a 95% confidence level with a 5% margin of error is widely accepted as a standard (Boston University, 2019). Using the Z-score value of 1.96 for 95%, $p = .5$, $q = .5$ and $e = .05$ the sample size n is calculated as:

$$n = \frac{z^2 pq}{e^2} = n = \frac{1.96^2 \times 0.5 \times 0.5}{0.05^2} = \frac{0.9604}{0.0025} = 384.16$$

Consequently, the Sample size = 385 for a very large or unknown population. Therefore, this research study required at least 385 responses to have the desired sample size.

3.2 Scope and Approach

A questionnaire was developed based on the framework created and research conducted in the literature review of data and storage security. The questionnaire was derived to allow individuals from various backgrounds, experiences, and technical knowledge to answer most of the questions in an appropriate amount of time. It was determined that the best method of eliciting the most objective and candid responses was to conduct an online survey that was 100% anonymous ensuring that the data collected could not be traced back to the person or organization.

The survey was shared on social media and through direct email requests for participation. The targeted participants' scope was strictly working adults at least 18 years of age

in the United States. As this was an anonymous survey, the inclusion and exclusion controls were addressed in this chapter's response validation section.

During the data collection, Excel was used to create a means of validating and cleaning the data to identify incomplete, irrelevant, or erroneous sample sets that may skew the results. The sections that follow describe the operations undertaken to confirm the data sample's totality used for analysis. Lastly, data analysis was performed in Excel and using Intellectus StatisticsTM (Intellectus) (Intellectus Statistics, 2020), an online statistical analysis tool.

3.3 Research Tool Development and Approval

The primary goal of the survey was to enable a means of answering the research questions (RQs) with real-world opinions, experiences, and knowledge from working adults. Both subjective and objective questions were created to analyze the opinions, expertise, and knowledge of working professionals from various industries. The survey included questions about the day-to-day usage and understanding of the operational security controls based on the seven principles introduced and demographic and Likert scale questions. The questions were classified into five categories (number of questions):

1. **Consent (1)** - agree or disagree at least 18 years old and consent terms
2. **Demographic (4)** - regional location, industry type, organization size, and role
3. **Polling (2)** – Response to data types and regulations associated with the organization
4. **Opinion (3)** – Likert scale opinion questions used to analyze the second and third RQ's
5. **Security Principle (11)** – question directly related to the seven data security model

The questionnaire mixed multiple-choice, single responses and selected all that apply, true and false, scaling questions, and hybrid questions. The hybrid questions gave the individual a list of answer responses and the freedom to fill in an answer without restriction. Additionally,

most questions gave the participant the option to opt-out of a question by answering, “I prefer not to answer” or “N/A” (Not Applicable).

The survey was hosted on SurveyMonkey and tested for understandability and readability by peers, committee members, and a small subset of outside individuals. The testing results allowed for correcting discrepancies in language or misunderstanding; these results were deleted. A copy of the final product can be seen in [Appendix B](#). The survey was approved by the DSU’s Institutional Review Board (IRB) ([Appendix C](#)).

3.4 Demographic Categorization

For Classification and qualification, several demographic questions were asked. The demographic areas were:

- **Regional:** *What region are you in?*
- **Industry:** *What is your organization’s industry or business type?*
- **Organization Size:** *How many employees are there in your organization?*
- **Working Role:** *Which area of your organization is your primary function or role?*

For the four demographic questions, the participant had a list of predefined responses and the option to fill in an answer. For question 2, region, only the response of the USA was qualifying. For question 3, industry, the response was classified into one of fourteen categories. [Appendix D](#) is a detailed list of all responses to question 3. For question 4, organization size, responses were classified as small, medium, or large, with [Appendix E](#) showing all responses. Lastly, for question 9, the work role was classified as technical, non-technical, educator, or leadership, with [Appendix F](#) showing these responses and how they were classified.

3.5 Response Validation

Validation of a survey involves determining the response meets specific requirements to reduce, as much as possible, disqualifying answers that may skew the results (Boussalis, 2016).

For a participant's response to be qualified, it required that it pass seven validation checks.

1. **Consent check:** Respondent must consent to terms of survey
2. **Regional check:** Respondent must reside in the United States
3. **Industry check:**
 - The respondent must be employed or identify with an industry
 - The respondent should answer based on an industry that can be classified
 - Example of disqualified: Homemaker, prefer not to answer (PNTA), or other responses that cannot be associated with an industry
4. **Work role check:**
 - The response must fit into one of the four role categories
 - Disqualified if retired without additional information, unemployed or PNTA
5. **Organization size check:** Disqualified if a response is unemployed or PNTA
6. **Number answers check:** The participant must have answered at least the first nine questions
7. **Timing check:** Respondent must have spent a minimal amount of time on the survey

Only responses that passed all seven of the validation checks qualified for analysis.

While checks three, four, and five may be considered subjective, all provided responses are open for review in the appendices. Check six and seven are strictly quantifiable based on how many questions were answered, and the start and end timestamp recorded from SurveyMonkey. While there were 21 numbered questions, the opinion Likert questions were treated as separate questions for analysis for a total of 38 responses. As mentioned, participants needed to answer at least nine questions, which translates to 25 responses; subsequently, answers to questions will be referenced by their response number when appropriate. The participant must take an average of

9 seconds per question; therefore, the total time needed to pass a completed survey is at least 5 minutes and 42 seconds. While this did not guarantee that the participant read and answered truthfully, it did remove responses where answers were clicked through without fully comprehending the questions. Table 12 shows that when a participant completed the survey, the average time spent was 8 minute and 37 seconds.

Table 12 Average time spent on the survey

Responses	# Responses	Average Time
Completed Survey	1,483	8m 37s
All Responses	2,255	6m 57s

3.6 Data Collection

The survey went live on July 4, 2020, following the IRB approval and ending on October 22, 2020, for a total of 16 weeks. A total of 2,255 individuals clicked on the survey link, with 95% agreeing to the first consent question. If consent was not given, the survey would end.

Figure 8 is a view of the distribution of qualified vs. disqualified responses.

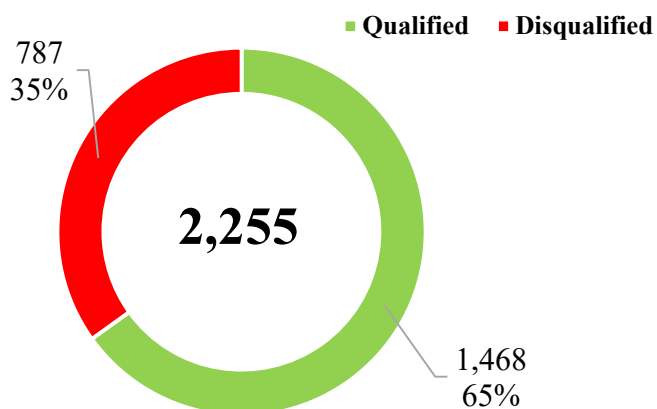


Figure 8 Distribution of qualified and disqualified responses

With 787 (~35%) disqualified, Table 13 shows, including those that did not consent, most, 89% (695), were disqualified by merely not answering a minimum number of questions.

Table 13 Results of validation check

Validation Check	Qualified	Disqualified	Total
Consent check	2,143	112	2255
Regional check	2,086	169	2255
Industry check	2,023	232	2255
Work tole check	1,542	713	2255
Organization size check	1,536	719	2255
Number answers check	1,560	695	2255
Timing check	1,631	624	2255

Quality assurance will decrease bias, improve precision, and improve the confidence that conclusions are supported by credible data (Riley, 2019). As described above, the first steps require removing non-responses, ensuring that the participants' assembled responses are qualified and represent the targeted population (Henry, 1998). As it does not change the results, all validation checks were shown for every participant in Table 13; however, there is overlap in disqualification criteria. The qualification phase results provided six valid responses, and ~94% were fully completed surveys. The sample size of 1,468 surpasses the minimum required sample size of 385. All data and analysis moving forward only address the qualified samples.

3.7 Data Processing Analysis

Survey analysis compiles meaningful answers from the raw data collected (Harrison, 2018). While this is a simple explanation, it is far more complicated in practice. As with any survey, responses that require binary or a numeric scaled response are relatively straight forward for evaluation. This is the case of the Likert questions (question 7, 8, and 21) and the true/false question number 19 (see [Appendix B](#)). However, when the question does not solicit measurable

value but rather a label or named response, it is nominal (Market Research Guy, 2020). Nominal questions and responses may be open to interpretation, but they can provide richer information (Glen, 2020).

As mentioned previously, the primary goal of the survey was to test the seven principles. This was accomplished by pulling security control mappings from Table 8 and developing questions to determine if the participant's control was implemented or selected. For example,

10. What are the requirements to access your organization's systems? (Select all that apply)

- Username/password
- Biometric identification (fingerprint, facial recognition, other)
- Two-factor authentication
- I don't know
- I prefer not to answer
- Other (please specify)

Examining the answer choices and cross-referencing the principle in Table 8, the principle tested is identified as *Authentication*. While there are over 100 controls identified in Table 8, it would make for an unmanageable survey to attempt to test for every control. For this reason, the survey questions used a small cross sampling of more comfortable to identify controls from each principle. [Appendix G](#) shows each question and how it is classified based on the purpose or security principle, making the questionnaire more inclusive. It should be noted that some questions served to test for more than one security control. Since most of these questions produce a nominal response, having a more extensive data sample is critical to reducing the outlier's impact and yielding a smaller error margin (Zamboni, 2018).

A scoring method was developed to convert nominal responses to a numeric value to determine if a security principle was passed. [Appendix H](#) shows the questions and possible security controls selected and the value-added or subtracted based on the scored principle. The

scoring rubric results were seven scale variables representing each security principle's score for every qualified survey response.

While the value assigned may seem subjective on its own, the determination for passing was based on a curve comparison between all participants. Scoring on the curve was done to avoid evaluation bias and evaluate the metric derived from the sample comparison rather than an anticipated response; if it is a good test, the result should form a bell curve showing a normal distribution (Roell, 2019). This scoring method is appropriate as the organizations' environment and security standards are unknown.

Summary statistics were calculated for Authentication score, Authorization score, Privacy score, Reliability score, Verification score, Recoverability score, and Accessibility score. When the skewness is greater than 2 in absolute value, the variable is asymmetrical about its *mean*. When the kurtosis is greater than or equal to 3, then the variable's distribution is markedly different from a normal distribution in its tendency to produce outliers (Westfall & Henning, 2013). The summary statistics can be found in Table 14.

Table 14 Summary Statistics Table for Interval and Ratio Variables (Intellectus Statistics, 2020)

Variable	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>	Min	Max	Skewness	Kurtosis
Authentication score	3.44	5.22	1468	0.14	-4.00	17.00	0.61	-0.14
Authorization score	4.60	4.51	1468	0.12	-4.00	13.00	0.00	-1.04
Privacy score	0.57	0.73	1468	0.02	-1.00	4.00	0.34	1.41
Reliability score	2.31	1.96	1468	0.05	-4.00	5.00	-0.80	0.22
Verification score	4.08	3.28	1468	0.09	-3.00	11.00	0.29	-0.64
Recoverability score	1.35	1.62	1468	0.04	-2.00	4.00	-0.72	-0.55

Accessibility score	4.39	2.12	1468	0.06	-2.00	10.00	-0.18	-0.24
---------------------	------	------	------	------	-------	-------	-------	-------

The results show that the skewness is less than 2 for all given principles scored, meaning that the values are relatively symmetrical about the *mean*, and Kurtosis is less than 3, showing no significant outliers. Therefore, this shows that each scoring is a normal distribution. [Appendix I](#) is a glossary of standard statistical terminology used throughout the research.

All scores were whole numbers; a passing value was determined to be the whole number left of the *mean* and greater than 64%. Table 15 shows the value calculation for the passing score of each principle.

Table 15 Calculating passing value criteria > 64% and < 70%

Variable Score	<i>M</i>	64 - 69%	Passing Score
Authentication score	3.44	2.3048	>2
Authorization score	4.6	3.082	>3
Privacy score	0.57	0.3819	>0
Reliability score	2.31	1.5477	>1
Verification score	4.08	2.7336	>2
Recoverability score	1.35	0.9045	>0
Accessibility score	4.39	2.3.03	>3

A Scatter (XY) graph was created in Excel. The *y-value* is the normal distribution, and the x-value is the principles scored. Figure 9 shows the bell curve created for the *Authentication* Scoring where scores in green would be passing.

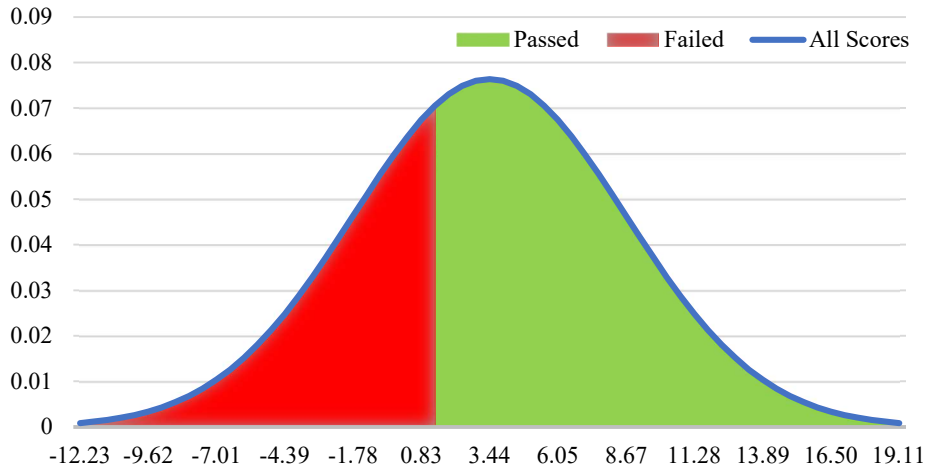
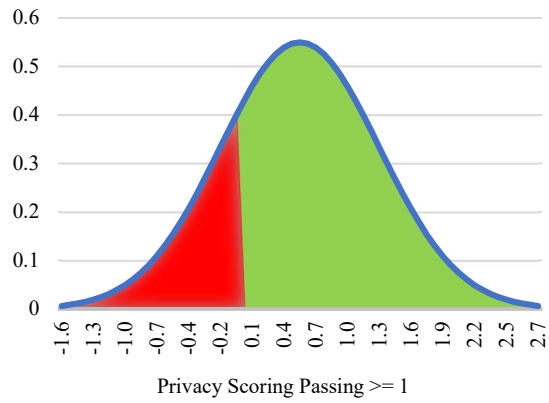
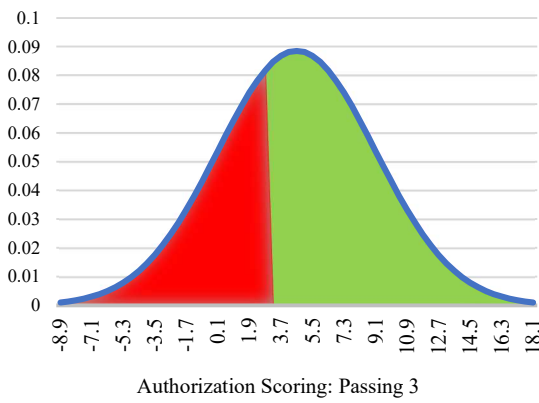


Figure 9 Authentication Scoring probability distribution passing score > 2

The graph shows the distribution scoring from lowest to highest values. The area under the curve in red was the cutoff for failures, and the green was passing. Below, Figure 10 are the bell curves for the remaining six principles.



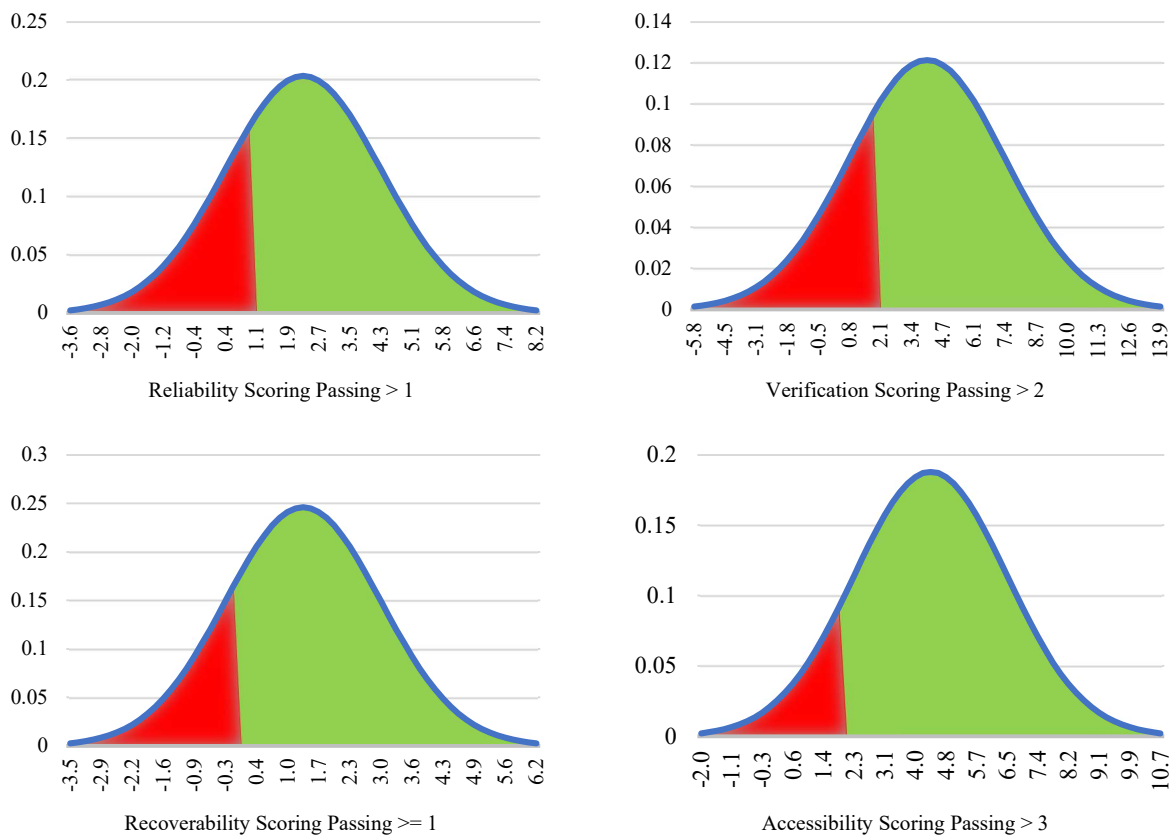


Figure 10 Scoring curves for other security principles

As seen in the above figures, each security principle's scoring shows bell curves validating a normal distribution among the samples scored. A new binary variable for a pass and fail was created for determining the results for all the security principles. This, in turn, resulted in an additional scale variable between 0 and 7 based on whether passing or failing was determined. The new variable will be referenced as *TotalPrinciplesPassed*.

3.8 Data Analysis Plan

Chapter 4 goes into details of the data analysis approach taken to address the three research questions. The techniques will include:

- A descriptive summary statistical view of the data

- Exploring differences between groups
- Predicting an outcome between variables
- Examining relationships between variables

The examination used summary metrics and richer distribution representations to detect the responses' essential features, effects, behaviors, or outliers to examine the research questions. The specific techniques that were used are described below, and a detailed glossary of terms can be found in [Appendix I](#).

3.8.1 Descriptive Statistics

Descriptive or summary statistics are typically used to describe or summarize the data. It is used as an exploratory method to examine the variables of interest, potentially before conducting inferential statistics on them. They provide summaries of the data and answer descriptive research questions (Intellectus Statistics, 2020).

3.8.2 Binary Logistic Regression

Binary logistic regression is used to examine the relationship between one or more independent (predictor) variables and a single dichotomous dependent (outcome) variable. This analysis aims to use the independent variables to estimate the probability that a case is a member of one group versus the other (Pituch & Stevens, 2015). The Binary Logistic Regression creates a linear combination of all the independent variables to predict the dependent variable's log-odds. In this analysis, the regression model's overall significance is tested by computing the χ^2 statistic used to compute the p-value (i.e., significance level). A significant overall model means that independent variables significantly predict the dependent variable (Menard, 2009). If the overall model is significant, the significance of each independent variable is assessed. An odds ratio (OR) is computed for each independent variable. It shows the extent that each independent

variable affects the probability that a case is a member of one outcome group versus the other. Conducting a Binary Logistic Regression requires that the dependent variable be dichotomous (i.e., there are only two possible outcomes). The observations must be independent of each other, and the relationship between the independent variables and the logit-transformed dependent variable must be linear (Intellectus Statistics, 2020).

3.8.3 Point Biserial Correlation

A Point Biserial correlation correlates with one dichotomous variable (a variable with only 2 unique values) and a continuous variable (Conover & Iman, 1981a). A correlation expresses the strength of linkage or co-occurrence between two variables in a single value between -1 and +1. This value that measures linkage strength is called the correlation coefficient, which is represented typically as the letter r (J. Cohen, 1988). A positive r value expresses a positive relationship between the two variables (the larger A becomes, the larger B becomes), while a negative r value indicates a negative relationship (the larger A becomes, the smaller B becomes). A correlation coefficient of zero indicates no relationship between the variables. However, correlations are limited to linear relationships between variables (Intellectus Statistics, 2020).

3.8.4 Spearman Correlation

A correlation expresses the strength of linkage or co-occurrence between two variables in a single value between -1 and +1 (J. Cohen, 1988). This value that measures linkage strength is called the correlation coefficient, which is represented typically as the letter r . A positive r value expresses a positive relationship between the two variables (the larger A becomes; the larger B becomes). In comparison, a negative r value indicates a negative relationship (the larger A

becomes, the smaller B becomes) (Conover & Iman, 1981b). A correlation coefficient of zero indicates no connection between the variables. Spearman rank correlation is a non-parametric test used to measure the degree of association between two variables. It was developed by Spearman; thus, it is called the Spearman rank correlation. Spearman rank correlation test does not make any assumptions about the distribution of the data and is the appropriate correlation analysis when the variables are measured on a scale that is at least ordinal level (J. Cohen, 1988).

3.8.5 Two-Tailed Independent Samples t-Test

The independent samples t-test is used to determine if there is a significant difference between two groups on a scale-level dependent variable. An independent samples t-test is the appropriate statistical test when the purpose of the research is to assess if differences exist on a continuous (interval/ratio) dependent variable by a dichotomous (2 groups) independent variable (Yeager, 2014). This test uses the difference between the two groups' average scores to compute the t statistic used with the df to compute the p-value (Intellectus Statistics, 2020). A significant result indicates that the observed test statistic would be unlikely under the null hypothesis. The independent samples t-test carries the assumptions of independence of observations, normality, and equality (or homogeneity) of variance (Razali & Wah, 2011).

3.8.6 Data sources

The data analysis was performed with Intellectus and Excel's aid using the calculated and qualified responses observed as scale and nominal variables. The calculated value for the number of principles passed would be referenced as TotalPrinciplesPassed. The responses from the following questions provided the primary independent and dependent variables.

Q6. "Which laws and regulations do you consider relevant to your organization?"

Respondents were given 11 regulations to select, PNTA, Unknown and fill in option

Q7. "How significant are the following consequences of unlawful, unauthorized, or accidental types of data incidents to your organization?"

(On the scale of 1 to 5, where 1 is "Not significant", 5 is "Very significant")

Reputation and brand damage – bad or embarrassing press

Reduced revenue or customer loss

Loss of trust on the part of interested parties

Litigation / legal proceedings

Deterioration of relations with employees'

Regulatory actions/sanctions or fines

Loss of competitive advantage (for example, due to loss of intellectual property)

Q21. "On a scale of 1 to 5, where 1 - "Strongly Disagree", 5 - "Strongly Agree", rate the following statements:"

My organization makes data security and privacy the highest priority

Figure 11 Questions that derive the primary independent and dependent variables

3.9 Summary

This chapter detailed the process and methodology for creating the primary survey research tool. It also calculated the necessary sample size needed based on the research goals and working adults' population. Throughout the research data collection process, the goal was to assure representative responses that would yield meaningful data that avoided bias. The results were 1,468 qualified responses, which significantly exceeded the required sample size. The data processing analysis section described the process used to analyze the seven principles to determine when a sample passed or failed. This process is a fundamental component that needs to convert the nominal responses into quantifiable measurements. Lastly, a brief outline of the statistical techniques was introduced to highlight the questions that will derive the dependent and independent variables needed to substantiate or refute the hypotheses. The next chapter will review the findings and analysis for this study.

CHAPTER 4: RESULTS AND ANALYSIS

As introduced in chapter 3, the primary research tool was an anonymous survey where the population was based on a random sample of working adults residing in the United States. The methodology and inclusion criteria were highly scrutinized to ensure the most objective results and ensure the best means of success to respond to the research questions without bias. This chapter thoroughly reviews the data collected, analyzed, and presented the result decisively using summary tables and clear visual graphs.

4.1 Demographics

Understanding the demographics information can provide greater insight into individuals' background characteristics participating in a survey (Dobronte, 2013). This can ensure that your audience is qualified and represents a good sample for the targeted population (Dobronte, 2013). However, asking too many descriptive or personal questions can make individuals uncomfortable, defensive, and compromise anonymity (Epstein, 2012). For this reason, it was decided to limit the demographic questions to just the most general to reassure participants. The first two questions, age consent followed by geographic location, were discussed in chapter 3 and were explicitly for qualification. The three other generic demographic categories were also discussed and used for qualification; however, job or work function, organization size, and industry types may be leveraged for future analysis. Figure 12 shows the summary graphs for the role category and the organization size for all 1468 qualified responses.

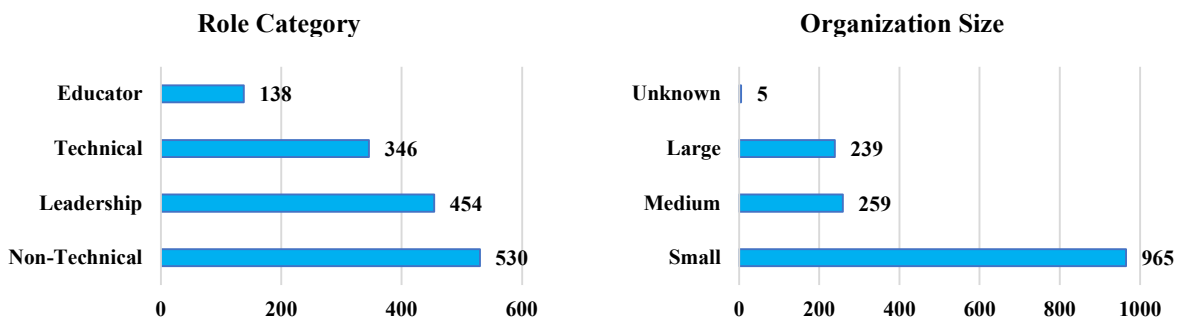


Figure 12 Number of qualified participants based on role category and organization size

The figures represent only qualified responses backed by the inclusion criteria explained in chapter 3 and supported by [Appendix E](#) and [Appendix F](#). Table 16 displays the industries cross-tabulated by role showing distribution, highlighting the top value in the category.

[Appendix D](#) shows the detailed list of industries and how they were classified.

Table 16 Qualified responses by Industry split by Role (green = highest %)

Industry	Frequency	Leadership	Non-Technical	Technical	Educator
Business Services	261	113 (43%)	90 (34%)	56 (21%)	2 (1%)
Education	201	20 (10%)	36 (18%)	30 (15%)	115 (57%)
Technology	162	30 (19%)	28 (17%)	103 (64%)	1 (1%)
Retail	149	57 (38%)	61 (41%)	31 (21%)	
Healthcare	136	28 (21%)	81 (60%)	23 (17%)	4 (3%)
Nonprofit	127	30 (24%)	67 (53%)	18 (14%)	12 (9%)
Manufacturing	105	48 (46%)	44 (42%)	12 (11%)	1 (1%)
Other	93	56 (60%)	27 (29%)	9 (10%)	1 (1%)
Financial	86	26 (30%)	47 (55%)	13 (15%)	
Government	62	18 (29%)	18 (29%)	24 (39%)	2 (3%)
Transportation	27	10 (37%)	7 (26%)	10 (37%)	
Utility	25	5 (20%)	8 (32%)	12 (48%)	
Real Estate	17	5 (29%)	10 (59%)	2 (12%)	
Media	17	8 (47%)	6 (35%)	3 (18%)	

Frequencies and percentages were calculated for Industry split by Role. The highest value for a given role by industry is highlighted in blue. For Leadership, the most frequently

observed Industry category was *for Business Services* ($n = 113$, 25%). For Technical, the most frequently observed category was *Technology* ($n = 103$, 30%). For Non-Technical, the most frequently observed category was *Business Services* ($n = 90$, 17%). For the Educator, the most frequently observed category was *Education* ($n = 115$, 83%).

Table 17 displays the industries cross-tabulated by organization size showing distribution, highlighting the top value in the category. [Appendix D](#) shows the detailed list of industries and how they were classified.

Table 17 Industry split by organization size (Green = highest)

Industry	Frequency	Small	Medium	Large
Business Services	261	229 (88%)	18 (7%)	14 (5%)
Education	201	44 (22%)	78 (39%)	79 (39%)
Technology	162	113 (70%)	16 (10%)	31 (19%)
Retail	149	118 (79%)	21 (14%)	10 (7%)
Healthcare	136	75 (55%)	30 (22%)	29 (21%)
Nonprofit	127	105 (83%)	21 (17%)	1 (1%)
Manufacturing	105	77 (73%)	22 (21%)	6 (6%)
Other	93	77 (83%)	13 (14%)	3 (3%)
Financial	86	60 (70%)	9 (10%)	16 (19%)
Government	62	14 (23%)	16 (26%)	32 (52%)
Transportation	27	12 (44%)	5 (19%)	10 (37%)
Utility	25	11 (44%)	7 (28%)	7 (28%)
Media	17	15 (88%)	1 (6%)	1 (6%)
Real Estate	17	15 (88%)	2 (12%)	

The most frequently observed Industry category was Business Services ($n = 229$, 24%) for Small. For Medium, the most frequently observed Industry category was Education ($n = 78$, 30%). For Large, the most frequently observed Industry category was Education ($n = 79$, 33%).

4.2 Data Types

In addition to the demographic information, participants were also asked what data types were relevant to their organization. Figure 13 provides the summary values.

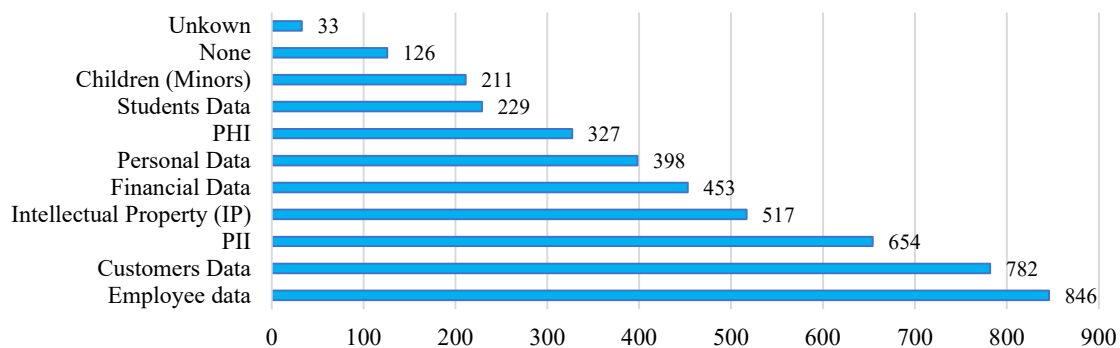


Figure 13 Number data types selected by qualified participants

The above graph showed that sensitive data, customers and employees, was recognized by over 50% of respondents. Simultaneously, data that may fall under regulated categories was only selected 20% to 40% by participants. Table 18 shows the Industry cross-referenced by the top 8 selected relevant data types.

Table 18 Industry split by top 8 selected data types (blue = Top 3)

Industry	Customers	Employee	PII	IP	PHI	Personal	Financial	Students
Business Services	158 (61%)	124 (48%)	113 (43%)	92 (35%)	52 (20%)	67 (26%)	77 (30%)	11 (4%)
Education	64 (32%)	140 (70%)	114 (57%)	82 (41%)	69 (34%)	89 (44%)	53 (26%)	158 (79%)
Technology	111 (69%)	92 (57%)	47 (29%)	91 (56%)	11 (7%)	22 (14%)	43 (27%)	9 (6%)
Retail	96 (64%)	72 (48%)	40 (27%)	31 (21%)	7 (5%)	13 (9%)	32 (21%)	1 (1%)
Healthcare	49 (36%)	72 (53%)	81 (60%)	33 (24%)	110 (81%)	76 (56%)	41 (30%)	13 (10%)
Nonprofit	50 (39%)	69 (54%)	52 (41%)	22 (17%)	16 (13%)	45 (35%)	28 (22%)	12 (9%)
Manufacturing	79 (75%)	72 (69%)	29 (28%)	56 (53%)	8 (8%)	7 (7%)	35 (33%)	0

Other	55 (59%)	65 (70%)	40 (43%)	26 (28%)	5 (5%)	12 (13%)	32 (34%)	1 (1%)
Financial	49 (57%)	47 (55%)	62 (72%)	24 (28%)	22 (26%)	32 (37%)	60 (70%)	12 (14%)
Government	24 (39%)	43 (69%)	44 (71%)	23 (37%)	22 (35%)	25 (40%)	25 (40%)	9 (15%)
Transportation	14 (52%)	21 (78%)	14 (52%)	17 (63%)	1 (4%)	5 (19%)	11 (41%)	1 (4%)
Utility	12 (48%)	16 (64%)	8 (32%)	11 (44%)	3 (12%)	3 (12%)	10 (40%)	0
Media	8 (47%)	5 (29%)	2 (12%)	5 (29%)	0	0	4 (24%)	1 (6%)
Real Estate	12 (71%)	7 (41%)	7 (41%)	3 (18%)	0	1 (6%)	1 (6%)	0

In the above table, the top 3 values for data types are selected for each industry. The takeaway from examining the data types is that individuals recognized multiple forms of data that may or may not be regulated for a given industry. For example, PHI (Protected Health Information) was selected by 327 out of 1468 participants, and only 34% (110 of 327) were selected by 136 that associated with healthcare.

4.3 Stored Data Security Principles Scoring Results

As discussed in chapter 3, the data security principles were scored based on the methodology of grouping the questions related to each principle and given a value. The passing value was determined on a bell curve and can be seen in chapter 3. The number that passed and failed is summarized in Figure 14.

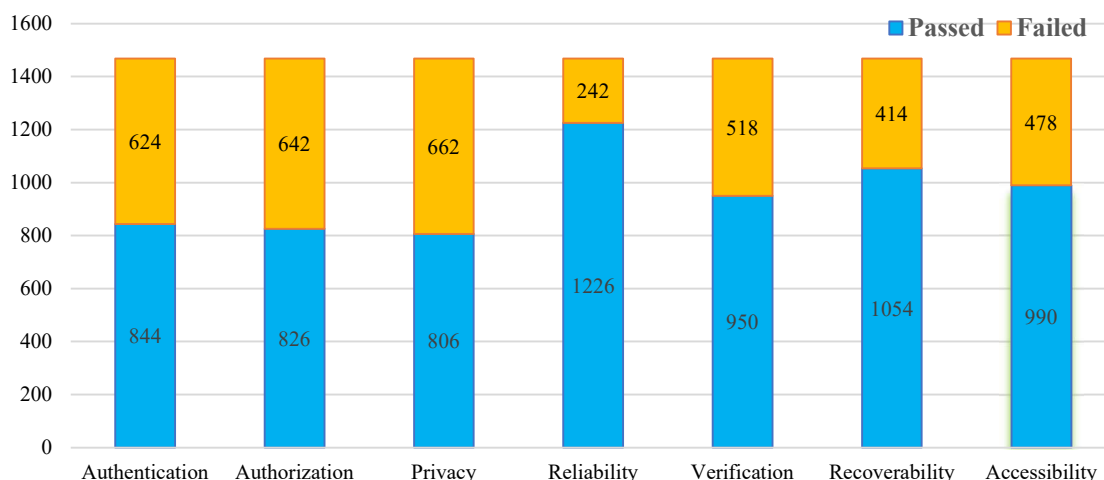


Figure 14 Pass/fail sums by Security Principle

The variable *TotalPrinciplesPassed* was calculated to count the number of failed principles for each of the 1468 responses. The value range was from 7 to 0 based on the passing principle tested. The summary of the results can be seen in Figure 15.

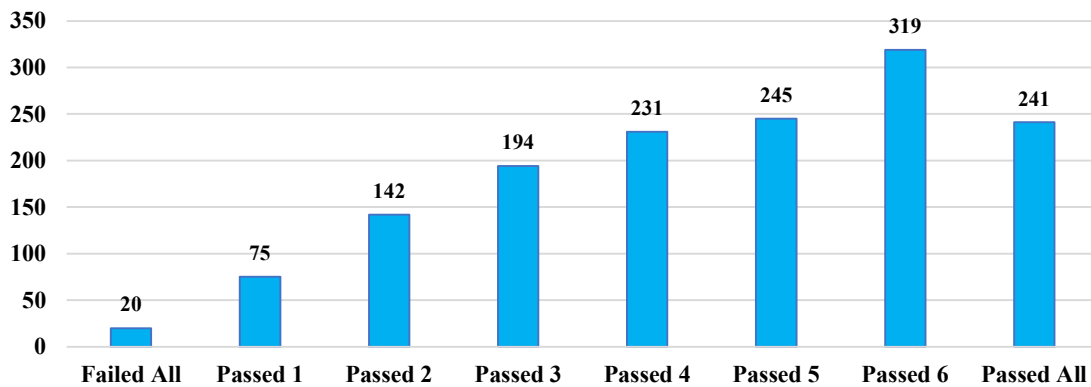


Figure 15 Total number of passed principles (*TotalPrinciplesPassed*)

4.4 Regulations

A baseline understanding of the regulations needs to be understood. [Appendix A](#) gives an overview of the 11 regulations and laws that participants could select as relevant to their organization and the option to respond by selecting “I don’t know” (*Unknown*) or *None*. Figure 16 shows how many of each regulation was selected.

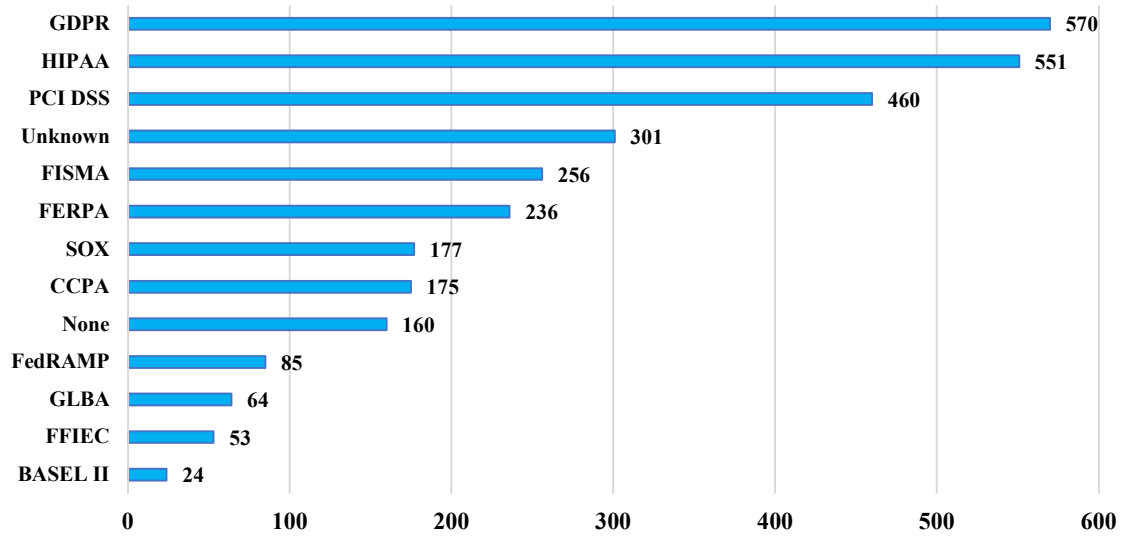


Figure 16 Regulation's choices selected

4.5 Research Question 1 Analysis

To begin the analysis, a brief review of the research question from chapter 1, the first research question (RQ₁) and supporting the null hypothesis (H₀1) and the alternative hypothesis (H₁1):

RQ1: If organizations comply with data privacy and security regulations, are they entirely securing stored data?

H₀1: Compliance regulations do not miss any of the seven (7) principles of stored data security.

H₁1: Compliance regulations miss at least one of the seven (7) principles of stored data security.

To accept or reject the hypotheses, it was required to examine the dependent variable *TotalPrinciplesPassed* to the independent regulations. The analysis was performed using several techniques described below.

4.5.1 Descriptive Analysis

Percentages were calculated for HIPAA, CCPA, GDPR, FERPA, FISMA, PCI_DSS, SOX, BASEL II, GLBA, FedRAMP, None and Unknown regulations have shown every value of the *TotalPrinciplesPassed* variable in Table 19.

Table 19 percentages of failed principles by each regulation response

Requisition	Frequency	Passed All	Failed > 0	Failed > 1	Failed > 2	Failed > 3	Failed > 4	Failed > 5	Failed All
HIPAA	570	131	438	293	199	116	58	23	3
PCI DSS	551	113	437	274	171	89	42	17	3
GDPR	459	104	355	244	167	99	51	17	0
FERPA	300	31	269	231	188	136	86	35	8
CCPA	256	80	175	99	61	33	17	6	1
SOX	236	59	176	102	61	33	15	7	2
FISMA	177	50	126	55	31	14	4	2	0
BASEL II	175	47	127	75	54	32	12	4	2
FFIEC	160	11	149	130	101	81	49	22	8
FedRAMP	85	31	53	25	15	11	5	2	0
GLBA	64	25	38	19	12	9	5	1	0
None Selected	53	19	33	22	14	6	4	1	0
Unknown	23	8	15	10	6	5	1	1	0

The results show the breakdown based on each regulation's values. The values were calculated using the *TotalPrinciplesPassed* variable for each regulation response option. This suggests a strong indication that each regulation is more likely to fail at least one principle.

Figure 17 shows how each regulation where the values are < 6.

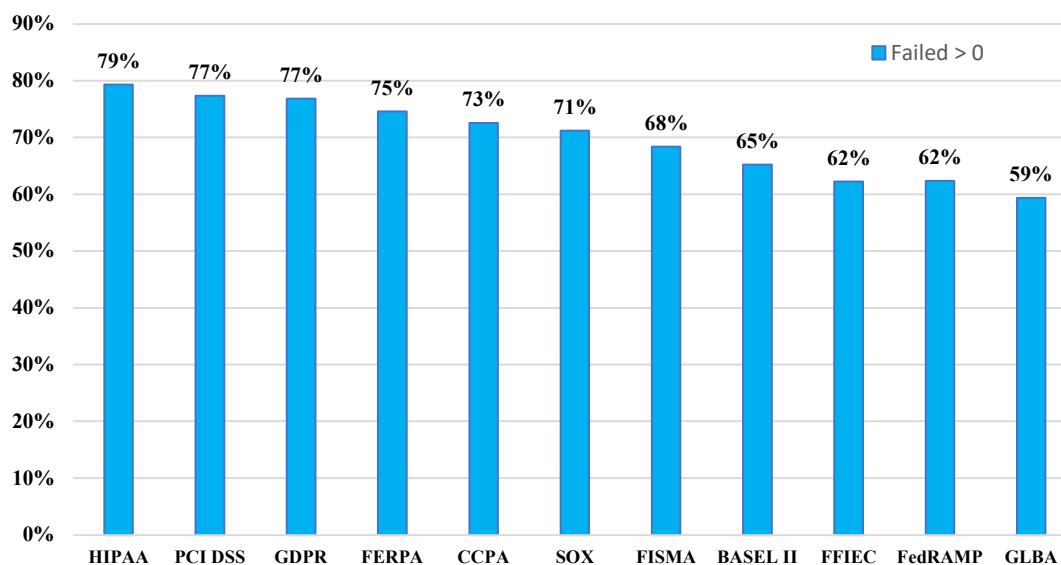


Figure 17 Regulations where TotalPrinciplesPassed < 7

The results suggest strong indications that the null hypothesis can be rejected. At face value, this implies that when analyzing the strengths of organizations' data security controls, at least one of the security principles examined will not pass ~74.5% of the time when regulations are required.

4.5.2 Binary Logistic Regression

It would be easy to accept the descriptive analysis results to justify rejecting the null hypothesis; however, this may conclude the relationship that regulations cause failure. For this reason, it is necessary to expand the analysis to examine if the results can be used to predict regulation selection.

For this reason, a Binary Logistic Regression was conducted using Intellectus to examine whether *TotalPrinciplesPassed* had a significant effect on the odds of observing a selection of at least one of the following regulation choices: GDPR, HIPAA, PCI DSS, FISMA, FERPA, SOX,

CCPA, FedRAMP, GLBA, FIEC, Basel II, Unknown or None. The reference category was the specific regulation not selected.

The complete Intellectus report can be found in [Appendix J](#), and Table 20 is a condensed version of all regulation results. Below is an overview of all the statistical terms in the table for reference:

- Unstandardized Beta (B): The slope of the predictor with the dependent variable
- Standard Error (SE): How much the beta coefficient (B) is expected to vary
- Chi-squared (χ^2): A test statistic based on the χ^2 distribution. Used with the degrees of freedom (df) to calculate a p-value
- p-value (p): The probability of obtaining the observed result if the null hypothesis is true. A result is usually considered significant if the p-value is $< .05$
- Odds Ratio (OR): Odds ratios compare the odds of two events. Odds ratios greater than 1 indicate the event is more likely to occur. Odds ratios less than 1 indicate the event is less likely to occur
- Confidence Interval (CI): An interval that is expected to contain the true value of a statistic in n% of repeated samples from the same probability distribution. n is based on the confidence level of the confidence interval
- Change: The approximately present age of increase (+) or decrease (-)

Table 20 Logistic Regression results with TotalPrinciplesPassed Approx. % change predicting the regulation

Variable	B	SE	χ^2	p	OR	95% CI	Change	Regulation
(Intercept)	-4.41	0.33	174.93	< .001	-	-	61%	SOX
TotalPrinciplesPassed	0.47	0.06	66.58	< .001	1.61	[1.43, 1.80]		
(Intercept)	-5.06	0.47	117.72	< .001	-	-	55%	FedRAMP
TotalPrinciplesPassed	0.44	0.08	29.87	< .001	1.55	[1.33, 1.82]		
(Intercept)	-5.3	0.53	99.44	< .001	-	-	53%	GLBA
TotalPrinciplesPassed	0.43	0.09	21.71	< .001	1.53	[1.28, 1.83]		
(Intercept)	-3.44	0.25	184.8	< .001	-	-	46%	FISMA
TotalPrinciplesPassed	0.38	0.05	68.7	< .001	1.46	[1.34, 1.60]		
(Intercept)	-4.93	0.52	88.71	< .001	-	-	38%	FIEC
TotalPrinciplesPassed	0.32	0.09	12.19	< .001	1.38	[1.15, 1.66]		
(Intercept)	-3.12	0.24	164	< .001	-	-	35%	FERPA
TotalPrinciplesPassed	0.3	0.04	44.62	< .001	1.35	[1.24, 1.47]		

(Intercept)	-1.85	0.17	124.79	< .001	-	-		
TotalPrinciplesPassed	0.28	0.03	77.73	< .001	1.33	[1.25, 1.42]	33%	HIPAA
(Intercept)	-5.48	0.74	54.67	< .001	-	-		
TotalPrinciplesPassed	0.27	0.13	4.1	0.043	1.31	[1.01, 1.70]	31%	Basel II
(Intercept)	-3.24	0.27	144.95	< .001	-	-		
TotalPrinciplesPassed	0.25	0.05	25.97	< .001	1.29	[1.17, 1.42]	29%	CCPA
(Intercept)	-1.52	0.16	93.59	< .001	-	-		
TotalPrinciplesPassed	0.23	0.03	54.39	< .001	1.26	[1.18, 1.34]	26%	GDPR
(Intercept)	-1.64	0.17	98.57	< .001	-	-		
TotalPrinciplesPassed	0.18	0.03	31.79	< .001	1.2	[1.13, 1.28]	20%	PCI DSS
(Intercept)	-0.21	0.16	1.72	0.189	-	-		
TotalPrinciplesPassed	-0.27	0.04	57.75	< .001	0.76	[0.71, 0.82]	-24%	Unknown
(Intercept)	-0.86	0.19	20.7	< .001	-	-		
TotalPrinciplesPassed	-0.3	0.04	43.81	< .001	0.74	[0.68, 0.81]	-26%	None

The regression coefficient for *TotalPrinciplesPassed* was significant with all results, indicating that for a one-unit increase in *TotalPrinciplesPassed*, the odds of observing any of the 11 regulations would increase, whereas, for *None* and *Unknown* would decrease. These results imply that regulations are not the causations of failure, but instead will improve the total number of principles passed.

4.5.3 Point Biserial Correlation Analysis

To fully understand the relationship between regulations and the total principles passed, it is necessary to measure association strength. Correlation analyses can be performed to express this strength of association in a single value, the correlation coefficient.

A Point Biserial correlation analysis was conducted between all the regulation responses and *TotalPrinciplesPassed*. A Point Biserial correlation is a special case of the Pearson correlation. Cohen's standard was used to evaluate the relationship's strength, where .1, .24, and .37 represent small, medium, and large effect sizes (J. Cohen, 1988). These effect size thresholds assume that both binary variable values are equally likely to occur (Rice & Harris, 2005).

The complete Intellectus report can be found in [Appendix K](#), and Table 21 is a condensed version of all regulation results. Below is an overview of all the statistical terms in the table for reference:

- Correlation Coefficient (r_{pb}): Ranges from -1 to 1, describes the strength of the relationship between the variables
- Confidence Interval (CI): An interval that is expected to contain the true value of a statistic in $n\%$ of repeated samples from the same probability distribution. n is based on the confidence level of the confidence interval
- p-value (p): The probability of obtaining the observed result if the null hypothesis is true. A result is usually considered significant if the p-value is $< .05$.

Table 21 Point Biserial correlations for Regulations and TotalPrinciplesPassed

Combination	r_{pb}	95% CI	p	Results
HIPAA - TotalPrinciplesPassed	-0.24	[-0.28, -0.19]	$< .001$	increases
FISMA - TotalPrinciplesPassed	-0.22	[-0.27, -0.18]	$< .001$	increases
SOX-Total - TotalPrinciplesPassed	-0.22	[-0.27, -0.18]	$< .001$	increases
GDPR - TotalPrinciplesPassed	-0.2	[-0.24, -0.15]	$< .001$	increases
FERPA - TotalPrinciplesPassed	-0.18	[-0.23, -0.13]	$< .001$	increases
FedRAMP - TotalPrinciplesPassed	-0.15	[-0.20, -0.10]	$< .001$	increases
PCI_DSS - TotalPrinciplesPassed	-0.15	[-0.20, -0.10]	$< .001$	increases
CCPA - TotalPrinciplesPassed	-0.14	[-0.19, -0.08]	$< .001$	increases
GLBA - TotalPrinciplesPassed	-0.13	[-0.18, -0.08]	$< .001$	increases
FFIEC - TotalPrinciplesPassed	-0.09	[-0.14, -0.04]	$< .001$	increases
BASEL II - TotalPrinciplesPassed	-0.05	[-0.10, -0.00]	0.039	increases
None - TotalPrinciplesPassed	0.18	[0.13, 0.23]	$< .001$	decreases
Reg. Unknown - TotalPrinciplesPassed	0.2	[0.15, 0.25]	$< .001$	decreases

The small effect size can be determined for each regulation above by looking at the correlation coefficient (r_{pb}), ranging between -0.24 and 0.2, indicating the small effect size.

When the results are determined to produce a small effect size, larger sample size is required.

Within the Intellectus tool, a power analysis was performed using G*Power. It was determined that the required sample size needed for a Point Biserial correlation with a small effect size

would be at least 779 (Intellectus Statistics, 2020). Since the sample size was 1468, the results of correlation were significant.

As seen above, the correlation coefficient values for the 11 regulations were negative, and for “none” and “unknown” were positive. This indicates that when the regulation was not selected, the *TotalPrinciplesPassed* would decrease. Selecting a regulation tends to be associated with an increase in the number of data security principles that passed. However, when the response was “none” or “unknown,” *TotalPrinciplesPassed* value would decrease.

By combining all regulations vs. combining “None” and “unknown,” the results suggest that while there is a ~74.5% chance of failing at least one principle, the percentage increased to ~90.9% when “None” or “Unknown” was selected. Figure 18 shows this comparison.

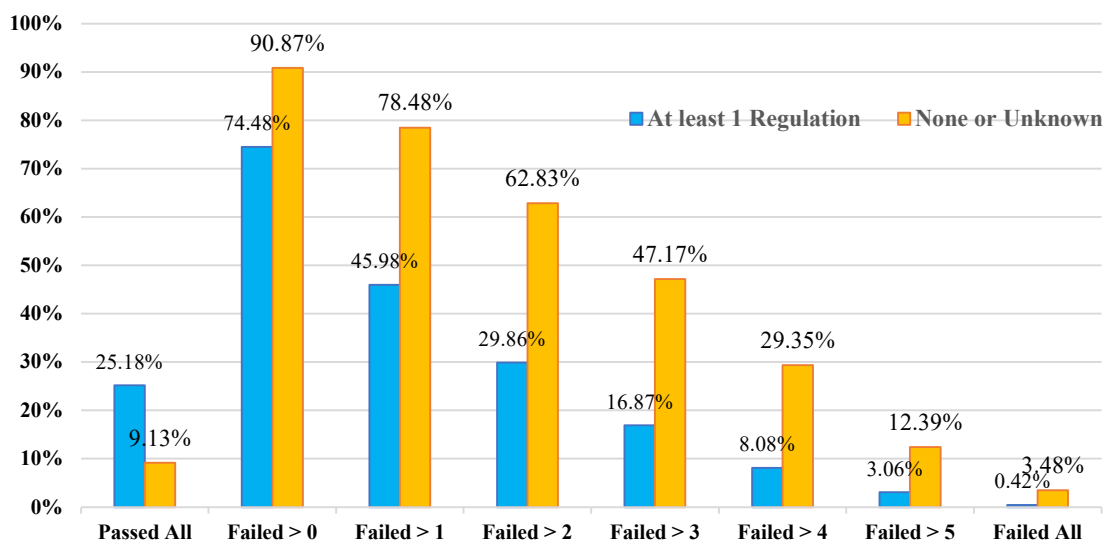


Figure 18 Present age of all regulations failing principle vs. None and Unknown combined

The results validate the binary logistic and descriptive analysis interpretation, supporting a strong indication that regulation does positively impact total principles passed. Nevertheless, the evidence suggests a gap still exists in a holistic view of data security.

4.5.4 Two-Tailed Independent Samples t-Test

Since the participants' option was to select or not select a regulation, a final analysis was conducted to determine independence between the two options. A Two-Tailed Independent Samples t-Test was conducted to examine whether the mean of *TotalPrinciplesPassed* was significantly different between all the regulations selected vs. if they were not selected categories.

The complete Intellectus report can be found in [Appendix L](#), and Table 22 is a condensed version of all regulation results. Below is an overview of all the statistical terms in the table for reference:

- Mean (*M*): The average value of a scale variable
- Standard Deviation (*SD*): The spread of the data around the mean of a scale variable
- t-statistic (*t*): Used with the *df* to determine the p-value
- p-value (*p*): The probability of obtaining the observed result if the null hypothesis is true. A result is usually considered significant if the p-value is $< .05$
- Cohen's *d* (*d*): Effect size for the t-test, determines the strength of the differences between the matched scores. The larger the effect size, the greater the differences in the matched pairs

Table 22 Two-Tailed Independent Samples t-Test for TotalPrinciplesPassed

Variable	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>d</i>
	HIPAA selected		HIPAA not selected				
TotalPrinciplesPassed	5.12	1.6	4.23	1.9	9.27	< .001	0.51
	FERPA selected		FERPA not selected				
TotalPrinciplesPassed	5.31	1.59	4.42	1.85	7.72	< .001	0.52
	GDPR selected		GDPR not selected				
TotalPrinciplesPassed	5.01	1.73	4.27	1.86	7.76	< .001	0.41
	CCPA selected		CCPA not selected				
TotalPrinciplesPassed	5.24	1.68	4.47	1.85	-5.63	< .001	0.44
	FISMA selected		FISMA not selected				
TotalPrinciplesPassed	5.46	1.58	4.37	1.84	9.71	< .001	0.64

	FedRAMP selected		FedRAMP not selected				
TotalPrinciplesPassed	5.68	1.55	4.49	1.84	6.71	< .001	0.7
	PCI DSS selected		PCI DSS not selected				
TotalPrinciplesPassed	4.97	1.73	4.38	1.87	5.92	< .001	0.33
	SOX selected		SOX not selected				
TotalPrinciplesPassed	5.68	1.3	4.41	1.85	11.49	< .001	0.79
	BASEL II selected		BASEL II not selected				
TotalPrinciplesPassed	5.35	1.75	4.55	1.84	2.06	0.039	0.44
	GLBA selected		GLBA not selected				
TotalPrinciplesPassed	5.67	1.62	4.51	1.84	5.51	< .001	0.67
	FFIEC selected		FIEC not selected				
TotalPrinciplesPassed	5.46	1.63	4.53	1.84	3.6	< .001	0.54
	Unknown selected		Unknown not selected				
TotalPrinciplesPassed	3.82	1.91	4.75	1.78	-7.61	< .001	0.5
	None selected		None not selected				
TotalPrinciplesPassed	3.62	1.9	4.68	1.8	-6.91	< .001	0.57

The results of the Two-Tailed Independent Samples t-Test were significant based on an alpha value of .05 for each case of the selected regulation indicating that there is a statistically significant difference in the *TotalPrinciplesPassed* by the categories of selecting a regulation vs. not selecting. This finding suggests the mean of *TotalPrinciplesPassed* increased when any regulation was selected; however, when “None” or “Unknown” was selected, *TotalPrinciplesPassed* would decrease. Figure 19 shows the results visual where the orange line is the “Not Selected” category and the blue bars are each regulation selection. When the bar is about the orange line, the mean value is more significant.

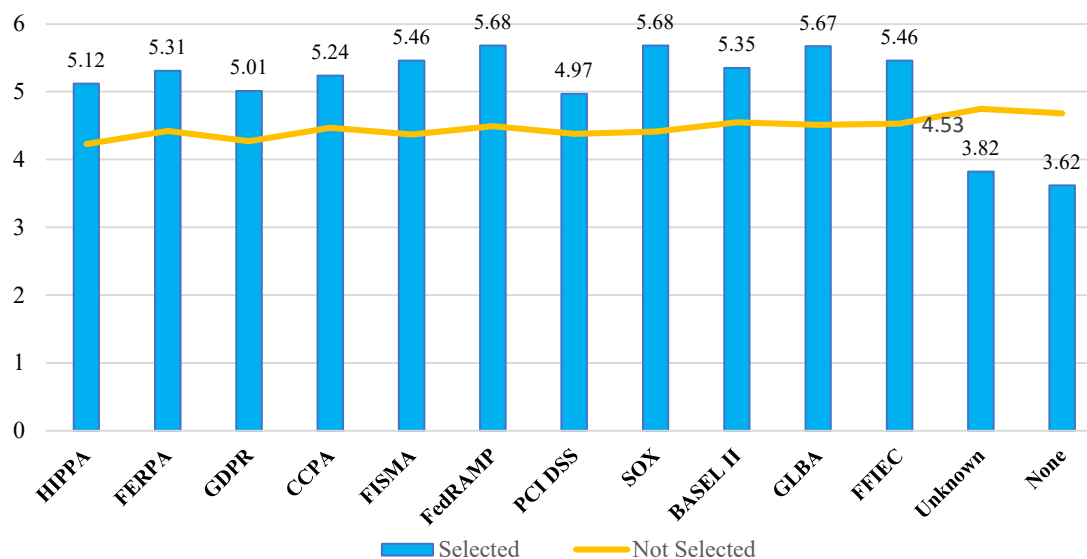


Figure 19 The mean of TotalPrinciplesPassed comparing regulations selected vs. not selected

Combining the results of all regulations and the effects of “*none* and *unknown*,” a clear picture is displayed of the strength of the selecting vs. not selecting. Table 23 shows the average value of the mean where the blue highlighted is the greater value of total principles passed.

Table 23 Averages of all means and total principles failed

	Average of the Means		Difference
	Selected	No Selected	
Regulation	5.36	4.42	+0.94
Principles Missed	1.64	2.58	-0.94
None or Unknown	3.72	4.72	-1.00
Principles Missed	3.28	2.29	+1.00

Lastly, the number of principles missed was also calculated by subtracting the combined average value from the max value for *TotalPrinciplesPassed* (7). This in turn shows how effective regulations are as an independent factor of predicting passing or failing all seven data security principles.

4.6 Research Question 2 and 3 Analysis

Chapter 1 also introduced the second and third research questions (RQ₂ and RQ₃) and supporting null hypotheses (H₀₂ and H₀₃) and the alternative hypotheses (H₁₂ and H₁₃):

RQ2: Does the social stigma of cyber incidents compel organizations to secure data?

H₀₂: Social stigma of data breaches is not more important to organizations than stored data security.

H₁₂: Social stigma of data breaches is more important to organizations than stored data security.

RQ3: Do data security laws, fines, and penalties compel organizations to implement stricter security controls for stored data?

H₀₃: Avoiding fines and penalties is not more important than data security to organizations.

H₁₃: Avoiding fines and penalties is more important than data security to organizations.

Since both research questions examine consequences related to data security incidents, the analysis can be combined. Within the survey questionnaire, questions 7 and 21 were designed to address RQ2 and RQ3.

Q7. “How significant are the following consequences of unlawful, unauthorized, or accidental types of data incidents to your organization?”

Q21. “On a scale of 1 to 5, where 1 -"Strongly Disagree", 5 - "Strongly Agree", rate the following statement: My organization makes data security and privacy the highest priority.”

Each of these Likert scale questions had the option to rank a response between 1 – 5 or “N/A.” Table 24 shows the eight scale questions from Q7 and Q21 by reference number from [Appendix G](#) along with the reference (variable) name.

Table 24 RQ2 and RQ3 variable names for analysis

Response #	Question	Variable Name
7	Reputation and brand damage – bad or embarrassing press	<i>ConseqReputation</i>
8	Reduced revenue or customer loss	<i>ConseqCustLoss</i>
9	Loss of trust on the part of interested parties	<i>ConseqLossTrust</i>
10	Litigation / legal proceedings	<i>ConseqLitigation</i>
11	Deterioration of relations with employees’	<i>ConseqEmpRelations</i>
12	Regulatory actions/sanctions or fines	<i>ConseqFines</i>
13	Loss of competitive advantage (for example, due to loss of intellectual property)	<i>ConseqLostCompAdv</i>
37	My organization makes data security and privacy the highest priority	<i>SecPriVal</i>

4.6.1 Descriptive Analysis

Summary statistics were calculated for *ConseqReputation*, *ConseqCustLoss*, *ConseqLossTrust*, *ConseqLitigation*, *ConseqEmpRelations*, *ConseqFines*, *ConseqLostCompAdv*, and *SecPriVal*. Table 25 shows the summary statistics for the variables. For a complete list of the standard statistical terms used, see [Appendix I](#).

Table 25 Consequences of incidence vs. importance of data security and privacy

Variable Name	<i>M</i>	<i>SD</i>	<i>n</i>	<i>SE_M</i>	Min	Max	Skewness	Kurtosis
ConseqLossTrust	4.17	1.24	1389	0.03	1	5	-1.41	0.83
ConseqReputation	4.03	1.33	1373	0.04	1	5	-1.2	0.14
SecPriVal	3.96	1.07	1391	0.03	1	5	-0.81	-0.09
ConseqCustLoss	3.81	1.41	1344	0.04	1	5	-0.85	-0.65
ConseqLitigation	3.8	1.41	1364	0.04	1	5	-0.82	-0.7
ConseqFines	3.56	1.51	1324	0.04	1	5	-0.55	-1.19
ConseqEmpRelations	3.43	1.46	1283	0.04	1	5	-0.44	-1.19
ConseqLostCompAdv	3.32	1.55	1264	0.04	1	5	-0.29	-1.43

The above table shows the mean value for each of the referenced values based on all available responses. The mean sorted the table from high to low, and the blue highlighted row shows where *SecPriVal* ranked. The results show that the following ranking of the various consequences vs. ranking of security and privacy:

1. Loss of trust on the part of interested parties
2. Reputation and brand damage – bad or embarrassing press
3. My organization makes data security and privacy the highest priority
4. Reduced revenue or customer loss
5. Litigation / legal proceedings
6. Regulatory actions/sanctions or fines
7. Deterioration of relations with employees'
8. Loss of competitive advantage (for example, due to loss of intellectual property)

The results suggest that data breaches' social stigma is more important than data security based solely on the mean values, which implies that the null hypothesis (H_02) can be rejected. Whereas avoiding fines and penalties mean values were lower than security and privacy, suggesting that we can accept the null hypothesis (H_03). However, this validation maybe flawed as each value is independent.

4.6.2 Spearman Correlation Analysis

A Spearman correlation analysis was conducted between the “consequence” variables and the *SecPriVal* variable. Cohen's standard was used to evaluate the strength of the relationship. Coefficients between .10 and .29 represent a small effect size, coefficients between .30 and .49 represent a moderate effect size, and coefficients above .50 indicate a large effect size (J. Cohen, 1988).

A Spearman correlation requires that the relationship between each pair of variables does not change direction (Conover & Iman, 1981a). This assumption is violated if the scatter plot points between any pair of variables appear to shift positively to negative or negative to positive relationships. Figure 20 presents the scatterplot for each of the correlations. A regression line has been added to assist the interpretation.

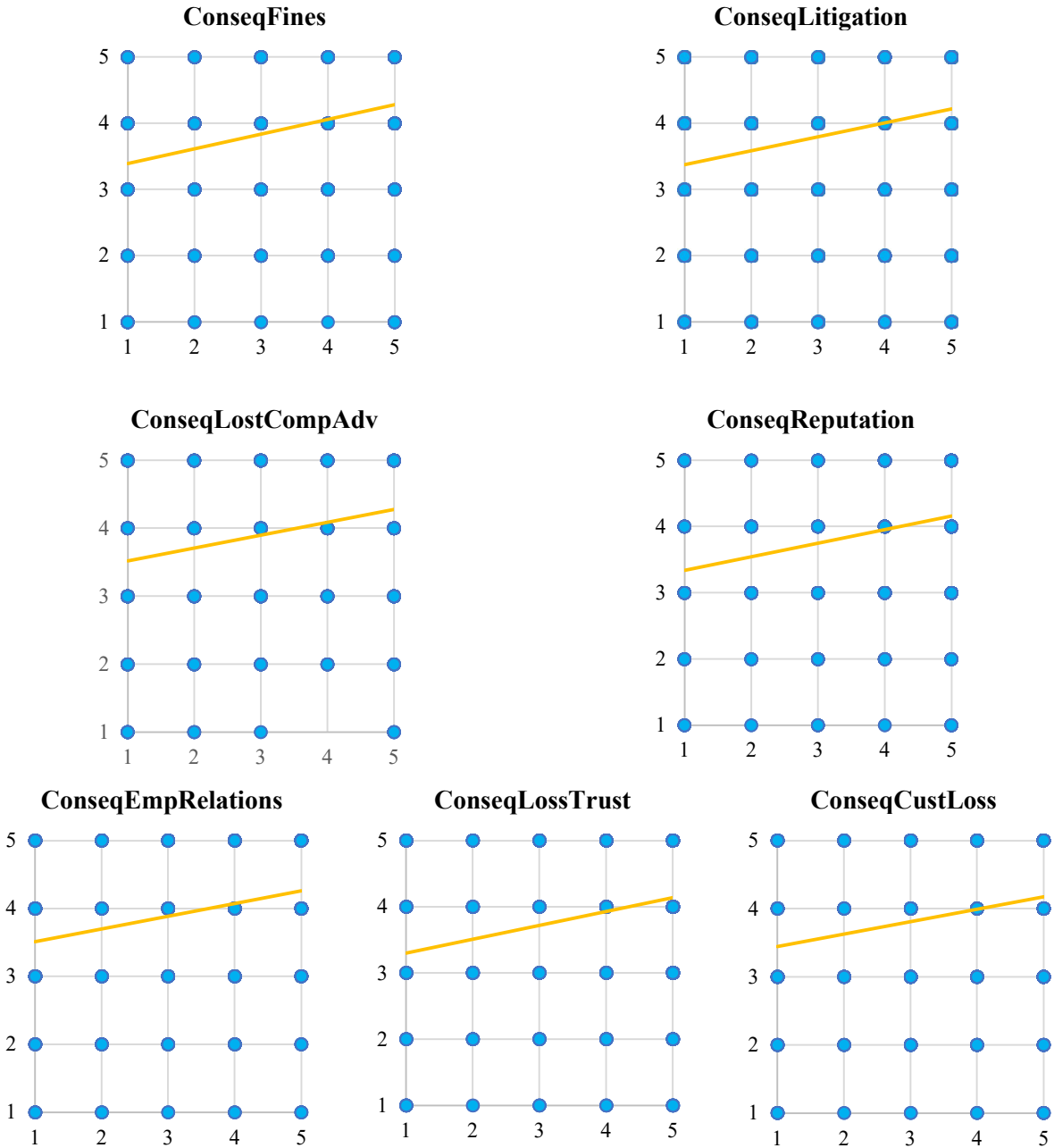


Figure 20 Scatterplots between each consequence with the regression line added

The complete Intellectus report can be found in [Appendix M](#), and Table 26 is a condensed version of all the consequences results. Below is an overview of all the statistical terms in the table for reference:

- Correlation Coefficient (r_s): Ranges from -1 to 1, describes the strength of the relationship between the variables
- Confidence Interval (CI): An interval that is expected to contain the true value of a statistic in n% of repeated samples from the same probability distribution. n is based on the confidence level of the confidence interval
- p-value (p): The probability of obtaining the observed result if the null hypothesis is true. A result is usually considered significant if the p-value is $< .05$.

Table 26 Spearman correlation results between consequences and SecPriVal

Combination	r_s	95% CI	p	Effect Size
ConseqFines - SecPriVal	0.34	[0.29, 0.38]	$< .001$	moderate
ConseqLitigation - SecPriVal	0.28	[0.23, 0.33]	$< .001$	Small
ConseqLostCompAdv - SecPriVal	0.28	[0.22, 0.33]	$< .001$	Small
ConseqReputation - SecPriVal	0.27	[0.22, 0.32]	$< .001$	Small
ConseqEmpRelations - SecPriVal	0.27	[0.22, 0.32]	$< .001$	Small
ConseqLossTrust - SecPriVal	0.26	[0.21, 0.31]	$< .001$	Small
ConseqCustLoss - SecPriVal	0.24	[0.19, 0.30]	$< .001$	Small

A significant positive correlation was observed between the “consequence” variables and *SecPriVal*. The table was sorted by the Correlation Coefficient (r_s) from high to low, showing *ConseqFines* and *SecPriVal* had a moderate effect size and all other “consequence” variables had a small effect size. All the correlations indicate that as the “consequence” variables increase, *SecPriVal* tends to increase.

Based on the correlation coefficient ranking, the result suggests the following order:

1. Regulatory actions/sanctions or fines
2. Litigation / legal proceedings
2. Loss of competitive advantage (for example, due to loss of intellectual property)
3. Reputation and brand damage – bad or embarrassing press
3. Deterioration of relations with employees’
4. Loss of trust on the part of interested parties

5. Reduced revenue or customer loss

Therefore, while the descriptive analysis results suggest only rejecting the null hypothesis **H₀₂**, the Spearman correlation results indicate that both null hypotheses, **H₀₂** and **H₀₃**, can be rejected. These results indicate a ranking of the seven consequences into five distinct levels. Where “Regulatory actions/sanctions or fines” had at least an 18% stronger relationship to data security and privacy than any other consequence.

Next, the *SecPriVal* was replaced with *TotalPrinciplesPassed* to determine if a relationship existed. Additionally, the Spearman correlation was performed between *SecPriVal* and *TotalPrinciplesPassed*. The scatter plot graphs for each correlation are not shown; however, the results were similar to Figure 20, showing a positive effect. The complete Intellectus report can be found in [Appendix N](#), and Table 27 is a condensed version of all the consequences results.

Table 27 Spearman correlation results between consequences and SecPriVal with TotalPrinciplesPassed

Combination	r_s	95% CI	p	Effect Size
SecPriVal - TotalPrinciplesPassed	0.37	[0.32, 0.41]	< .001	moderate
ConseqLitigation - TotalPrinciplesPassed	0.19	[0.14, 0.24]	< .001	small
ConseqLossTrust - TotalPrinciplesPassed	0.17	[0.12, 0.22]	< .001	small
ConseqFines - TotalPrinciplesPassed	0.17	[0.12, 0.22]	< .001	small
ConseqReputation - TotalPrinciplesPassed	0.16	[0.11, 0.21]	< .001	small
ConseqCustLoss - TotalPrinciplesPassed	0.15	[0.10, 0.20]	< .001	small
ConseqLostCompAdv - TotalPrinciplesPassed	0.15	[0.10, 0.20]	< .001	small
ConseqEmpRelations - TotalPrinciplesPassed	0.13	[0.08, 0.19]	< .001	small

A significant positive correlation was observed between the variables and *TotalPrinciplesPassed*. The table was sorted by the Correlation Coefficient (r_s) from high to low, showing SecPriVal and *TotalPrinciplesPassed* had a moderate effect size. All other “consequence” variables had a small effect size. The results suggested that when the

consequence value increases, the number of principles may increase slightly. However, the significance of the ranking of “My organization makes data security and privacy the highest priority” was at least 195% greater than any consequence for improving the number of principles passed.

4.7 Summary

This chapter summarized the survey's qualified data to address the three research questions. The summary statistics for the demographic and data types were examined to show the breakdown of participants' backgrounds and show the awareness and knowledge of the categories of data utilized in the modern workforce. The descriptive examination was also conducted of the seven data security principles' calculated scoring and the eleven compliance regulations. Multiple statistical techniques were utilized to determine strengths or weaknesses. In the next and final chapter, the findings and interpretations will be discussed as a conclusion of this study.

CHAPTER 5: CONCLUSION

The primary aim of this dissertation was to determine how effective regulations are for securing data. Secondly, it analyzed if social, legal, and financial repercussions had influenced stored data security awareness. A literature review revealed seven principles of stored data security used as a framework for identifying physical, administrative, and technical controls used as a reference for the development of an exploratory questionnaire. This chapter contains an interpretation of the findings as they relate to answering the research questions and will include a summary of the limitations, discussion of findings, and recommendations for future research. Portions of this research were peer reviewed and appeared in *Advances in Intelligent Systems and Computing Volume 1271*:

Goodman, H., & Rowland, P. (2020). Deficiencies of Compliancy for Data and Storage Isolating the CIA Triad Components to Identify Gaps to Security. *National Cyber Summit (NCS) Research Track 2020, 1271*, 170.

5.1 Limitations

As mentioned in chapter 1, the research was limited only to individuals that were willing to participate. Typically, only "pilot sample" studies are conducted for data-collection testing, preliminary information for planning, or future research when performing this research type (University of Baltimore, 2004). While the number of participants was prolific, many more answers would be needed to narrow the analysis to each specific industry. This limited the analysis to a more general interpretation of the sample population. While this does not invalidate the findings, it suggests that the results are backed solely on the sample. Since the survey was anonymous, the interpretation of individual principles scoring was based on a bell curve.

Furthermore, the number of questions was limited to a small sampling of quickly answered questions that people would be comfortable, knowledgeable, and encouraged to respond to. This limited the data to a subset selection of security controls for each stored security principle. For this reason, many higher technical security control questions were not asked as only individuals with direct knowledge of specific security information would know. For example, asking how data protection is addressed or questions on physical stored data security controls may have caused confusion or frustration if participants didn't understand or know the answer. Therefore, it is possible that the organizations addressed issues without the knowledge of participants limiting the accuracy of the responses.

5.2 Discussion of Findings

This research study combined the interpretation of well-known security controls, concerns, and opinions to develop an online survey. The survey was hosted on SurveyMonkey to instill confidence as a familiar platform with industry professionals and leverage the secure collector capabilities to meet the requirements for a broad random sample of participants. 1,468 qualified responses were filtered from the 2,255 total contributors using the techniques discussed in chapter 3. A complete breakdown of participant demographics was shown in chapter 4 for review. For analysis purposes, several statistical methods were utilized for validation and interpretation of the data.

The research analysis was performed using the following five statistical techniques:

1. Descriptive
2. Binary Logistic Regression
3. Point Biserial correlation
4. Two-Tailed Independent Samples t-Test
5. Spearman correlation

Chapter 3 presented an overview description of each method, and the results were analyzed in chapter 4, and the full analysis reports were presented in the appendices. The descriptive summary statistics were utilized for a basic summarized observable view of the results, whereas the other techniques gave a more advanced data analysis.

To determine if a security principle passed or failed, questions and answers were classified and assigned a value from the rubric in [Appendix H](#). Then, the values were tallied for each of the seven principles, and a grade of passing or failing was determined independently and impartially based on a bell curve described in chapter 3. The sum of passing grades was calculated, and a new variable called *TotalPrinciplesPassed* was created with a scaled value range of [0 – 7]. This variable was the baseline measurement for comparison.

5.2.1 Research Question 1

Participants were asked which data security regulations and laws were relevant to their organization. The question gave the respondents a choice of eleven data protection regulations, the option to write in a response, and answer unknown or none. The responses were collected into 13 independent binary variables where the value was either “selected” or “not selected.” The name of each variable was the regulation acronym, *None* or *Unknown*. These 13 binary variables, along with the scale variable *TotalPrinciplesPassed*, provided the statistical analysis data to answer RQ1.

The analysis for RQ1 was accomplished using four of the statistical methods. First, the summary data was analyzed, showing how each of the 13 regulation variables was selected and how often *TotalPrinciplesPassed* would fail. The results were shown in Table 19, with Figure 17 showing that ~74.5% of the time, at least one of the seven data security principles would fail

when one or more regulation was selected. This could imply that regulations cause at least one failure; however, a relationship or link existing between each regulation variable and *TotalPrinciplesPassed* would need to be determined.

Two tests were performed to determine if a relationship or link exists between each regulation variable and *TotalPrinciplesPassed*. First, a Binary Logistic Regression was performed between *TotalPrinciplesPassed* and each regulation variable as defined in chapter 3. The Binary Logistic Regression tested if a relationship existed between the independent predictor variable (*TotalPrinciplesPassed*) and a dichotomous dependent variable (each regulation variable). A summary of the results was shown in Table 20, and the full report was presented in [Appendix J](#). The results showed a 20 to 61% increase to the value of *TotalPrinciplesPassed* when the regulation was selected, whereas when *Unknown* or *None* was selected, *TotalPrinciplesPassed* would decrease by 24 or 26%.

Next, a Point Biserial correlation was used to explore the associated strength of the relationships between *TotalPrinciplesPassed* and each regulation variable. The association is determined by calculating the correlation coefficient represented by a value from -1 to 1. A positive value represents a positive relationship, and a negative value represents a negative relationship. The closer the correlation coefficient is to 1, the stronger the positive connection, whereas the closer to -1, the stronger the negative connection. A full explanation can be found in chapter 4, which showed that when a regulation was selected, the number of principles passed would increase; however, when respondents choose “no regulations” (*None*) or “not sure” (*Unknown*), the number of principles passed would decrease. Table 28 shows each regulation response and the top three principles that failed.

Table 28 Top 3 stored data security principles that failed for each selected regulation variable

Regulation Response	Top Principle Failed	Second Most Failed	Third Most Failed
Not Applicable	Authorization (73%)	Authentication (70%)	Privacy (60%)
Unknown	Authentication (64%)	Authorization (63%)	Verification (58%)
HIPAA	Privacy (54%)	Authentication (36%)	Authorization (36%)
PCI DSS	Authorization (50%)	Privacy (49%)	Authentication (45%)
GDPR	Privacy (57%)	Authorization (45%)	Authentication (40%)
FERPA	Privacy (67%)	Verification (35%)	Recoverability (30%)
CCPA	Privacy (63%)	Authentication (39%)	Recoverability (32%)
SOX	Privacy (62%)	Recoverability (25%)	Authorization (23%)
FISMA	Privacy (55%)	Authorization (34%)	Authentication (32%)
FFIEC	Privacy (55%)	Authorization (42%)	Authentication (39%)
FedRAMP	Privacy (68%)	Recoverability (32%)	Verification (26%)
GLBA	Privacy (55%)	Authentication (39%)	Authorization (37%)
Basel II	Privacy (67%)	Recoverability (40%)	Authorization (33%) Verification (33%)

While it became apparent that a principle was likely to fail, the above table showed the top 3 expected principles based on the regulation response. Table 29 ranked the seven stored data security principles based on which failed the most to least.

Table 29 Ranking of most likely failed stored data security principle

Principle	Average Failure	Ranking
Privacy	662 (45%)	1
Authorization	642 (44%)	2
Authentication	624 (43%)	3
Verification	518 (35%)	4
Accessibility	478 (33%)	5
Recoverability	414 (28%)	6
Reliability	242 (16%)	7

Finally, the Two-Tailed Independent Samples t-Test was performed to examine whether the mean of *TotalPrinciplesPassed* was significantly different between each regulation's options for validating independence. The results indicated that the mean value for *TotalPrinciplesPassed* would increase when a "regulation" was selected. In contrast, when the variables "None" or "Unknown" was chosen, the mean for *TotalPrinciplesPassed* would decrease. Figure 21 shows

the mean value of *TotalPrinciplesPassed* comparison between “Selected” and “Not Selected” for “Regulations” vs “None or Unknown.”

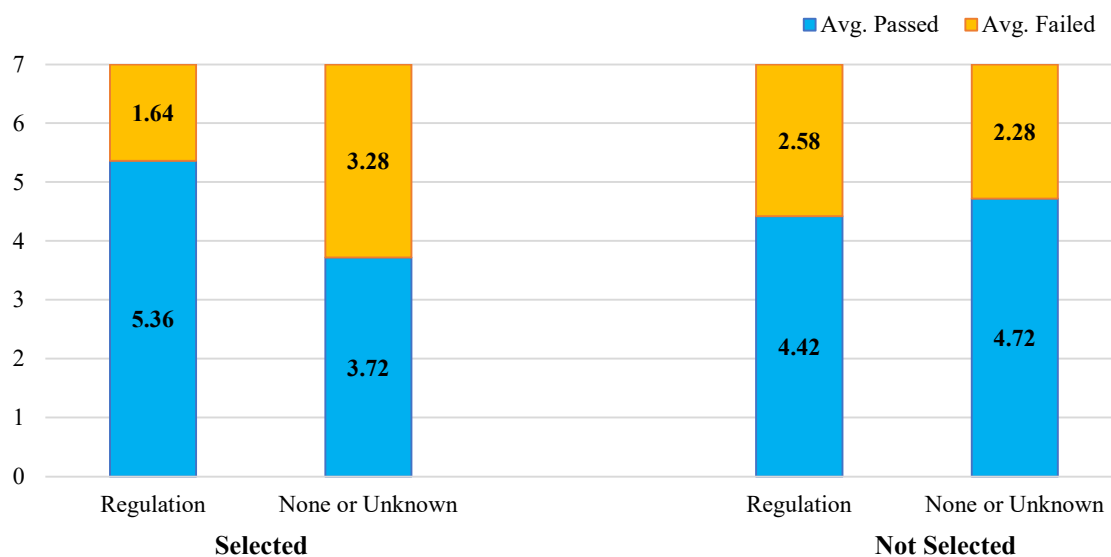


Figure 21 The difference for *TotalPrinciplesPassed* mean values regulation variables

The combined results in Figure 21 showed that when regulations are relevant to a respondent, at least 1.64 stored data security principles will fail. Nevertheless, the results showed a 200% increase of failed principles when "None or Unknown" were chosen. Additionally, the mean value for "Not Selected" showed only a .30 gap between the combined variables signifying no real difference.

5.2.2 Research Question 2 and 3

The analysis for RQ2 and RQ3 was performed simultaneously using the same response questions discussed in chapter 4 (Q7 and Q21). Q7 asked participants to rank the importance of seven consequences of a severe or damaging data incident or event. Each consequence was

given an independent value from "1 to 5", where "5" was most important. Q21 was the last survey question, which asked for a rank from "1 to 5" of how important data security and privacy are to their organization, where "5" was most important. It is noteworthy that the questions were separated for independence. The scaled value for consequences was given a variable name with a preface of "*Conseq*" followed by abridging descriptive name (see Table 24). The scaled ranking of "data security and privacy" was given the variable name *SecPriVal*.

The analysis was performed using a Spearman Correlation and examining the descriptive summary statistics. The summary statistics showed the observable mean value for each of the "consequence" variables and how it compared to the *SecPriVal* variable. The Figure 22 graph shows the summary from chapter 4.

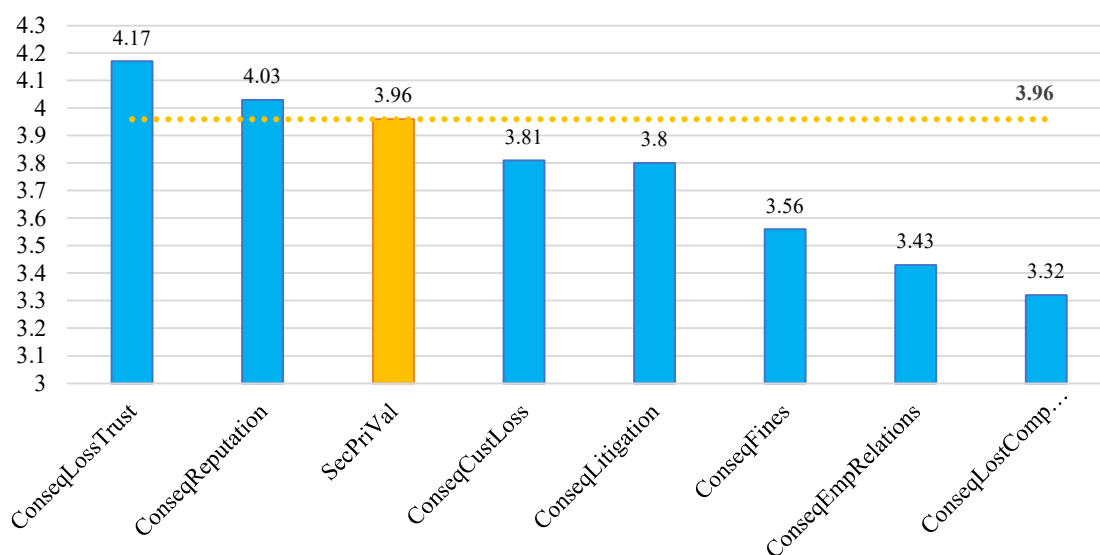


Figure 22 Ranking of consequences vs. Stored Data Security

The above graph showed the average value for responses to each of the consequences and *SecPriVal* (importance for stored data security and privacy). The line represents the *mean* for

SecPriVal, where consequences above the dotted line were ranked higher, and those below were of lesser concern. The results indicated the top trepidations were "loss of trust" followed by "Reputation and brand damage," which implies more significant concern for social stigma over other negative consequences. To further explore the data, two separate Spearman Correlations were conducted.

As discussed in chapter 3 and 4, the Spearman correlation test measures the strength and direction between two ranked variables. First, an analysis was performed between each of the "consequence" variables and *SecPriVal*. The tests were performed independently between each "consequence" variable and *SecPriVal*, which showed that as any consequence variable increased, so did *SecPriVal* indicating a positive relationship. Table 26 in chapter 4 showed the result ranked by the correlation coefficient. Figure 23 shows the results for each "consequence" variable correlated by *SecPriVal* sorted by the correlation coefficient.

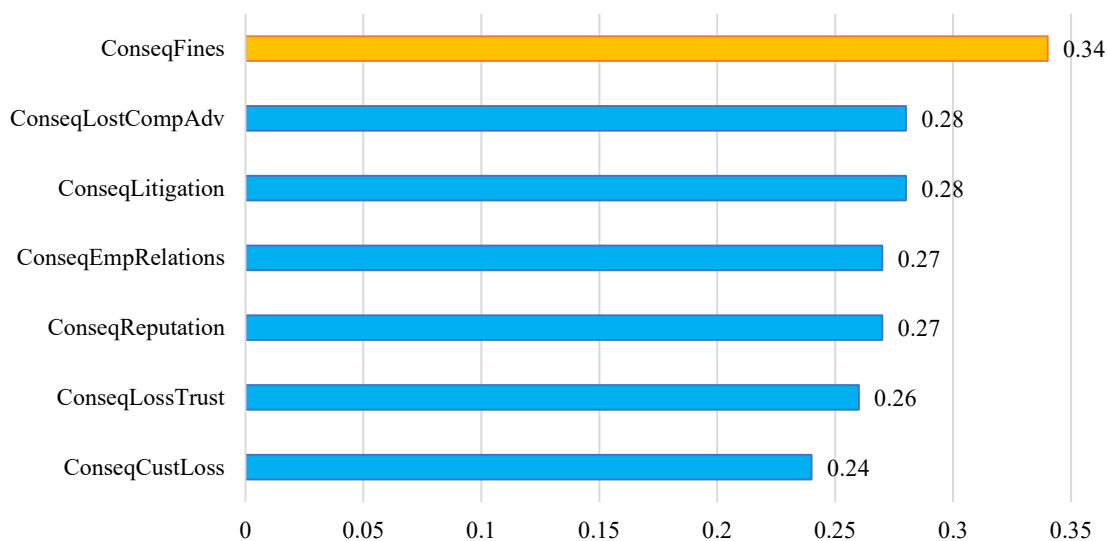


Figure 23 Spearman's ranking for each consequence variable and *SecPriVal*

The results of the Spearman correlation showed a different interpretation compared to the descriptive analysis. As each "consequence" variable increased, so did *SecPriVal*; however, in the investigation, the variables that affected *SecPriVal* were not the variables with the greater mean values. The figure shows that "*Regulatory actions/sanctions or fines*" had a more significant effect on *SecPriVal* than any other consequence, contradicting a straightforward interpretation of the descriptive analysis. Conversely, while *ConseqFines* showed more significant effect, each "consequence" variable showed a direct positive relationship.

Finally, a Spearman correlation was conducted on the participants' rankings to measure the effect on stored data security. All "consequence" variables and *SecPriVal* were evaluated with *TotalPrinciplesPassed* independently to measure the level of association. Table 26, in chapter 4, showed each evaluation's combined results, and the ranking of the correlation coefficients can be seen in Figure 24.

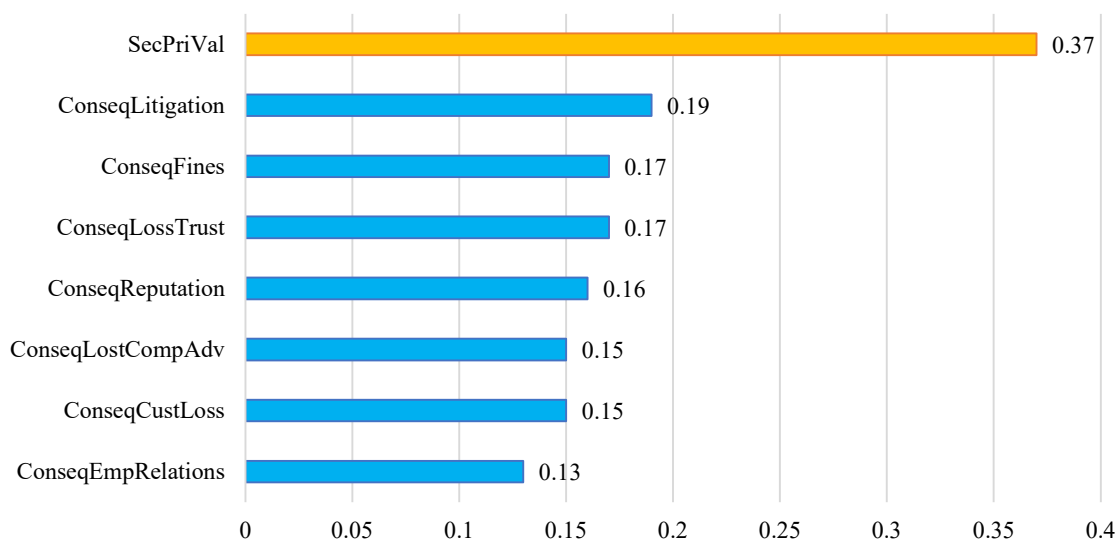


Figure 24 Ranking of between ConseqFines and *SecPriVal* for *TotalPrinciplesPassed*

The above chart showed that *SecPriVal* had a 195% - 285% greater effect on improving the value of *TotalPrinciplesPassed* over any of the "consequence" variables. Statistically, each consequence did have an impact on increasing the number of stored data principles. Again, as seen in Figure 23, reputation concerns did not reflect the descriptive analysis results.

5.3 Summary and Future Works

The research conducted in this dissertation was performed from 2019 – 2021 to understand better organizations' awareness, requirements, concerns, and strategies for stored data security. A detailed literature review on data and storage security, 11 United States data protection regulations, and various consequences and imposed penalties was conducted. The coalesced investigation identified seven principles for data security and facilitated creating the research survey to respond to the research questions. To answer the three research questions, null and alternative hypotheses were formulated to evaluate the sample data derived from survey participants.

2,143 individuals consented to participate, and 1,468 were qualified after meeting all the validation requirements. After the sample data was collected and processed, four statistical tests were performed to analyze the first research question, and three were conducted for the second and third research questions. Two work streams were performed where the first was performed on research question 1, and the second was completed on research questions 2 and 3 as they both originated from the same survey questions. Below is the highlighted summarized response for each research question (RQ).

RQ1. If organizations comply with data privacy and security regulations, are they entirely securing stored data?

- Regulations failed at least one principle 74.5% of the time

- When regulations weren't selected, one principle would fail 90.9% of the time
 - The results indicated that regulations positively improve stored data regulations compared to when regulations weren't selected
- RQ2. Does the social stigma of cyber incidents compel organizations to secure data?
- The top concern was the loss of trust followed by reputation and brand damage, ranking higher than concerns for data security and privacy
 - The results indicated that the social stigma of cyber incidences do compel organizations to secure data; however, to a lesser extent than legal, fines and loss of competitive advantage
- RQ3. Do data security laws, fines, and penalties compel organizations to implement stricter security controls for stored data?
- Regulatory actions, sanctions, or fines were less critical than four of the seven other consequences and not as significant as concerns for data security and privacy
 - Regulatory actions, sanctions, or fines showed a moderate effect size on how important participants ranked data security and privacy
 - The results suggested laws, fines and penalties do compel organizations to implement stricter security controls for stored data

From the analysis performed for research question 1, it was observed on all four tests that there was a strong indication that at least one of the stored data principles would fail. However, it was also observed that when regulations were not selected, the number of principles that failed would double, indicating regulations had a positive effect on the number of principles passed. The statistical evidence showed that the top principle most likely to fail was privacy for ten of the eleven regulations and the principle least likely to fail was Reliability.

In exploring research questions 2 and 3, the three statistical tests' results weren't as decisive as the previous analysis. By examining the average rankings for all consequences

compared to how important data security and privacy were, only trust and brand reputation were more significant. However, the correlation analysis exhibited a small positive effect for all consequences except for "Regulatory actions/sanctions or fines," whose positive impact was moderate, indicating an 18 – 29% greater influence over the other results. The results suggested that all consequences positively affected data security and privacy, contradicting a direct hierarchical comparison of the mean values.

Finally, the rankings were correlated to the seven data security principles' results, which revealed that the higher-ranked value of the importance of data security and privacy had a 195% - 285% greater positive effect on the total number of principles that passed compared to the consequences. While the results indicated consequence values positively impacted the seven data security principles, the two analyses' combined view seemed to imply a more substantial indirect effect. The greater the ranking for consequences stimulated a higher importance value for data security and privacy, where data security and privacy promoted an increase in the number of data security principles that passed.

The following future research recommendation will benefit and expand on the efforts of this study. While no direct evidence can be concluded based on this study as to why privacy was 1.5 – 4 times more likely to fail than any other principle, it does support the need for future research. As noted, this study was limited to anonymous participation from individuals working within the United States. If the survey was from an identifiable sample set, the scoring could have been based on best practice security controls or industry standards. Meaning, if the sample were identifiable, the data interpretation could have been empirically validated by well-known or verified or requirements. For example, if the respondent's specific organization name was confirmed, the applicable regulations could have been cross-referenced to the selected response.

Therefore, it is recommended to conduct an informed confidential consensual targeted survey for a specific organization or regulated industry sector.

In conclusion, the defining moment for any organization's data security posture arises when an unforeseen incident occurs. Data protection regulations only define the requirements, guidelines, enforcement, and penalty for non-compliance, whereas the impact of an incident affects data security when consequences lead to more significant concern for data security and privacy. Mitigation, response, and correction are the security controls' objectives.

The seven stored data security principles formulated and discussed above can help organizations understand the gaps, limitations, and corrective actions needed to advance stored data security.

REFERENCES

- 451 Group. (2020). 451 Group. In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=451_Group&oldid=957619151
- Achary, N. (2019, June 12). The Importance of Data Privacy. Retrieved November 11, 2020, from Medium website: <https://medium.com/@neelachary/the-importance-of-data-privacy-39c6676eeb58>
- Agarwal, A., & Agarwal, A. (2011). *The Security Risks Associated with Cloud Computing. 1*, International Journal of Computer Applications in Engineering Sciences.
- Agrawal, A. (2016, February). Configuring NIC Teaming in Windows Server 2016. Retrieved November 30, 2020, from Study.com website: <https://study.com/academy/lesson/configuring-nic-teaming-in-windows-server-2012-r2.html>
- Alexandra Twin. (2020, July 28). How a Non-Disclosure Agreement (NDA) Works [Non-Disclosure Agreement (NDA)]. Retrieved November 29, 2020, from Investopedia website: <https://www.investopedia.com/terms/n/nda.asp>
- Apple. (2020, November 20). Hide My Email for Sign in with Apple. Retrieved December 1, 2020, from Apple Support website: <https://support.apple.com/en-us/HT210425>
- ASTHO. (2014, October 16). Comparison of FERPA and HIPAA Privacy Rule | State Public Health | ASTHO. Retrieved December 7, 2020, from <https://www.astho.org/programs/preparedness/public-health-emergency-law/public-health-and-schools-toolkit/comparison-of-ferpa-and-hipaa-privacy-rule/>
- Baig, A. (2020, March 23). Where Does Encryption Fit in Privacy Regulations? Retrieved August 22, 2020, from CMSWire.com website: <https://www.cmswire.com/information-management/where-does-encryption-fit-in-privacy-regulations/>
- Bansal, S. (2020, October 14). What is the Federal Information Security Management Act (FISMA?). Retrieved December 7, 2020, from <https://securityscorecard.com/blog/what-is-the-federal-information-security-management-act>

- BCI. (2018, February 9). Cyber-attack top business threat for third year running. Retrieved December 24, 2020, from <https://www.thebci.org/news/cyber-attack-top-business-threat-for-third-year-running.html>
- Beal, V. (n.d.). What is Structured Data? Webopedia Definition. Retrieved December 13, 2019, from https://www.webopedia.com/TERM/S/structured_data.html
- Becerra, X. (2018, October 15). California Consumer Privacy Act (CCPA) [CA Government]. Retrieved November 11, 2020, from State of California—Department of Justice—Office of the Attorney General website: <https://oag.ca.gov/privacy/ccpa>
- BIS. (2004). *Basel II: Revised international capital framework*. Retrieved from <https://www.bis.org/publ/bcbsca.htm>
- BlockApps. (Dec2017). How Blockchain Will Disrupt Data Storage. Retrieved December 15, 2019, from BlockApps website: <https://blockapps.net/blockchain-disrupt-data-storage/>
- Blum, R., & Singh, R. (2017). *Google—Site Reliability Engineering*. Sebastopol, CA: O'Reilly. Retrieved from <https://landing.google.com/sre/sre-book/chapters/data-integrity/>
- Bollinger, J., Enright, B., & Valite, M. (2015). *Crafting the InfoSec Playbook: Security Monitoring and Incident Response Master Plan* (1st ed.). O'O'Reilly`.
- Boston University. (2019, October 9). Confidence Intervals for Sample Size Less Than 30. Retrieved December 10, 2020, from <https://sphweb.bumc.bu.edu/otlt/MPH-Modules/PH717-QuantCore/PH717-Module6-RandomError/PH717-Module6-RandomError11.html>
- Boussalis, C. (2016). Basic Survey Theory and Design. *Harvard Law School*, 26.
- Bowen, P., Hash, J., & Wilson, M. (2006). Information Security Handbook: A Guide for Managers. *NIST Special Publication 800-100*, 178.
- Brandão, L., Davidson, M., Mouha, N., & Vassilev, A. (2019). ITL BULLETIN FOR APRIL 2019 TIME TO STANDARDIZE THRESHOLD SCHEMES FOR CRYPTOGRAPHIC PRIMITIVES. *NIST*, 6.
- Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019, April). Thematic analysis—The University of Auckland [Education]. Retrieved March 31, 2020, from Thematic analysis |

- a reflexive approach website: <https://www.psych.auckland.ac.nz/en/about/thematic-analysis.html>
- Brooks, R. (2019, November 5). Data Privacy Trends, Issues and Concerns. Retrieved April 3, 2020, from <https://blog.netwrix.com/> website: <https://blog.netwrix.com/2019/11/05/data-privacy-trends-issues-and-concerns-for-2020/>
- Burnette, M. (2018, March 19). Why and How to Disclose Data Breaches. Retrieved November 30, 2020, from LBMC Family of Companies website: <https://www.lbmc.com/blog/why-and-how-to-disclose-data-breaches/>
- Butler, K. R. B., McLaughlin, S. E., & McDaniel, P. D. (2007). Non-volatile Memory and Disks: Avenues for Policy Architectures. *Proceedings of the 2007 ACM Workshop on Computer Security Architecture*, 77–84. New York, NY, USA: ACM.
<https://doi.org/10.1145/1314466.1314479>
- Butler, K. R. B., McLaughlin, S., & McDaniel, P. D. (2008). Rootkit-resistant disks. *Proceedings of the 15th ACM Conference on Computer and Communications Security - CCS '08*, 403. Alexandria, Virginia, USA: ACM Press. <https://doi.org/10.1145/1455770.1455821>
- Butler, S. (2018, August 20). The Pros and Cons of Data Encryption | TechNadu. Retrieved August 23, 2020, from Technadu website: <https://www.technadu.com/pros-and-cons-of-data-encryption/38599/>
- Canadian Centre for Cyber Security. (2020, November 18). Canadian Centre for Cyber Security. Retrieved March 4, 2021, from Canadian Centre for Cyber Security website: <https://cyber.gc.ca/en/>
- Carlson, M., & Espy, J. (2017, January). *IP-Based Drive Management Specification*. SNIA. Retrieved from https://www.snia.org/sites/default/files/technical_work/IPdrive/IPBasedDriveMgmtSpecV1.0.pdf
- Carnegie Mellon University. (2018, May 23). Guidelines for Data Classification—Information Security Office—Computing Services—Carnegie Mellon University. Retrieved November 30, 2020, from <http://www.cmu.edu/iso/governance/guidelines/data-classification.html>

- Carnegie Mellon University. (2020). Guidelines for Data Protection—Application Security. Retrieved August 31, 2020, from <http://www.cmu.edu/iso/governance/guidelines/data-protection/application-security.html>
- Caulfield, J. (2019, September 6). How to Do Thematic Analysis | A Step-by-Step Guide & Examples. Retrieved April 1, 2020, from Scribbr website: <https://www.scribbr.com/methodology/thematic-analysis/>
- CDC. (2019, February 21). Health Insurance Portability and Accountability Act of 1996 (HIPAA) | CDC. Retrieved December 7, 2020, from <https://www.cdc.gov/phlp/publications/topic/hipaa.html>
- CFI. (2020, September 20). Basel II - Overview, Three Pillars, Components. Retrieved December 6, 2020, from Corporate Finance Institute website: <https://corporatefinanceinstitute.com/resources/knowledge/finance/basel-ii/>
- Chang, Z., & Hao, Y. (2009, October). *The research of disaster recovery about the network storage system base on "Safety Zone."* IEEE. Retrieved from <https://ieeexplore.ieee.org/document/5361098>
- Chen, M., & Zadok, E. (2019). Kurma: Secure geo-distributed multi-cloud storage gateways. *Proceedings of the 12th ACM International Conference on Systems and Storage - SYSTOR '19*, 109–120. Haifa, Israel: ACM. <https://doi.org/10.1145/3319647.3325830>
- CISA. (2019, December 26). Federal Information Security Modernization Act | CISA. Retrieved December 7, 2020, from <https://www.cisa.gov/federal-information-security-modernization-act>
- Cisco. (n.d.). What Is a Data Center? Retrieved December 13, 2019, from Cisco website: <https://www.cisco.com/c/en/us/solutions/data-center-virtualization/what-is-a-data-center.html>
- Classified information. (2020). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Classified_information&oldid=991252117

- Clearswift. (2019, December 17). The consequences of a data breach: Why fines are just the tip of the iceberg. Retrieved December 24, 2020, from Clearswift website: <https://www.clearswift.com/blog/2019/12/17/data-breach-consequences-beyond-fines>
- Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). Wiley.
- Cohen, A. (2019, October). Twenty Years of Successful Co-Regulation Under COPPA: A Model for Fostering Consumer Privacy. Retrieved November 11, 2020, from BBBPrograms website: https://bbbnp-bbbp-stf-use1-01.s3.amazonaws.com/docs/default-source/whitepapers/bbb-np-report---20-years-of-coppa-self-regulation---10-15-2019.pdf?sfvrsn=387d0185_2
- Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). St. Paul, MN: West Publishing Company.
- Collins, B. (2020, April 18). What is NIC Teaming? | My Virtual Journey. Retrieved November 30, 2020, from <https://www.myvirtualjourney.com/what-is-nic-teaming-and-why-we-use-it/>
- Conover, W. J., & Iman, R. L. (1981a). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician*, 124–129.
- Conover, W. J., & Iman, R. L. (1981b). Rank Transformations as a Bridge between Parametric and Nonparametric Statistics. *The American Statistician*, 35(3), 124–129. <https://doi.org/10.1080/00031305.1981.10479327>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed). Thousand Oaks: SAGE Publications.
- CyberInsureOne. (2019, April 19). Cybersecurity Laws and Penalties. Retrieved December 24, 2020, from CyberInsureOne website: <https://cyberinsureone.com/laws-penalties/>
- Daniel, E., & Vasanthi, N. A. (2019). LDAP: A lightweight deduplication and auditing protocol for secure data storage in cloud environment. *Cluster Computing*, 22(1), 1247–1258. <https://doi.org/10.1007/s10586-017-1382-6>

- DataGuidance. (2018, November 28). *Comparing privacy laws: GDPR v. CCPA*. DataGuidance. Retrieved from https://fpf.org/wp-content/uploads/2018/11/GDPR_CCPA_Comparison-Guide.pdf
- Dell EMC. (2018, December). *Dell EMC UnityTM Family Security Configuration Guide*. Dell EMC. Retrieved from <https://www.dell.com/ro-ro/collaterals/unauth/technical-guides-support-information/products/storage/docu69321.pdf>
- Dennis G. (2020, December). RTO vs. RPO: Two Means Toward the Same End. Retrieved December 6, 2020, from MSP360 Blog website: <https://www.msp360.com/resources/blog/rto-vs-rpo-difference/>
- Dharma, R., Venugopal, V., Sake, S., & Dinh, V. (2013, April). *Building Secure SANs*. EMC. Retrieved from <https://www.slideshare.net/emcacademics/h8082-buildingsecurasanstb-11556458>
- Dobronte, A. (2013, August 13). The importance of socio-demographics in online surveys. Retrieved December 30, 2020, from CheckMarket website: <https://www.checkmarket.com/blog/socio-demographics-online-surveys/>
- Domingo, A. I. S., & Villar, N. D. (2018). Self-regulation in data protection. *BBVA Research*, 4.
- Dunham, R. (2018, September 19). FedRAMP Compliance: What is it? Requirements, Process & More! Retrieved December 6, 2020, from Linford & Company LLP website: <https://linfordco.com/blog/fedramp-compliance/>
- Dutch, M. (2010, June). *A Data Protection Taxonomy*. SNIA. Retrieved from https://www.snia.org/sites/default/files/A_Data_Protection_Taxonomy_V51.pdf
- Eliezerov, R. (2020, January 9). Consumer Privacy and Data Protection Trends for 2020 [Industry news media]. Retrieved November 11, 2020, from The Future of Customer Engagement and Experience website: <https://www.the-future-of-commerce.com/2020/01/09/data-protection-trends-2020/>
- ENISA. (2019, January 28). *ENISA Threat Landscape Report 2018 15 Top Cyberthreats and Trends*. ENISA. Retrieved from <https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018>

- Epstein, L. (2012, December 18). Sensitive survey questions: What to do, what not to do. Retrieved December 30, 2020, from SurveyMonkey website: <https://www.surveymonkey.com/curiosity/sensitive-topics-methodology/>
- European Commission. (n.d.). What are Data Protection Authorities (DPAs)? [Text]. Retrieved November 11, 2020, from European Commission—European Commission website: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-are-data-protection-authorities-dpas_en
- Fearn, N. (2018, March 28). GDPR: Consumers demand more data privacy from the IoT. Retrieved November 9, 2020, from Internet of Business website: <https://internetofbusiness.com/consumers-demand-more-data-privacy-from-the-iot-economist-report/>
- FedRAMP PMO. (2018). *FedRAMP Continuous Monitoring Performance Management Guide. Version 2.1*, 12.
- FFIEC. (1996, December 18). FFIEC Home Page. Retrieved December 8, 2020, from Welcome to the Federal Financial Institutions Examination Council's (FFIEC) Web Site. website: <https://www.ffiec.gov/>
- Frankenfield, J. (2020, November 11). General Data Protection Regulation (GDPR). Retrieved December 7, 2020, from Investopedia website: <https://www.investopedia.com/terms/g/general-data-protection-regulation-gdpr.asp>
- Fruhlinger, J. (2020, February 10). The CIA triad: Definition, components and examples [Industry source]. Retrieved August 17, 2020, from CSO Online website: <https://www.csoonline.com/article/3519908/the-cia-triad-definition-components-and-examples.html>
- FTC. (2002, July 2). How To Comply with the Privacy of Consumer Financial Information Rule of the Gramm-Leach-Bliley Act. Retrieved December 8, 2020, from Federal Trade Commission website: <https://www.ftc.gov/tips-advice/business-center/guidance/how-comply-privacy-consumer-financial-information-rule-gramm>
- Fuller, C. S. (2019). Is the market for digital privacy a failure? *Public Choice*, 180(3), 353–381. <https://doi.org/10.1007/s11127-019-00642-2>

- Fuxi, G., & Yang, W. (2015). *Data Storage at the Nanoscale* (1st ed.). emny Stanford Publishing. Retrieved from <http://www.panstanford.com/pdf/9789814613200fm.pdf>
- Garun, N. (2019, May 29). How to set up two-factor authentication on all your online accounts [Industry]. Retrieved August 17, 2020, from The Verge website: <https://www.theverge.com/2017/6/17/15772142/how-to-set-up-two-factor-authentication>
- GDPR.EU. (2018a, July 11). What are the GDPR Fines? Retrieved December 7, 2020, from GDPR.eu website: <https://gdpr.eu/fines/>
- GDPR.EU. (2018b, November 5). Everything you need to know about the “Right to be forgotten.” Retrieved March 4, 2021, from GDPR.eu website: <https://gdpr.eu/right-to-be-forgotten/>
- Glen, S. (2020, September 23). Nominal Ordinal Interval Ratio & Cardinal: Examples. Retrieved December 17, 2020, from Statistics How To website: <https://www.statisticshowto.com/nominal-ordinal-interval-ratio/>
- Goodman, H., & Rowland, P. (2020). Deficiencies of Compliancy for Data and Storage Isolating the CIA Triad Components to Identify Gaps to Security. *National Cyber Summit (NCS) Research Track 2020, 1271*, 170.
- Gordan, J. (2019). *Practical Data Security (Unicom Applied Information Technology) First Edition Edition*.
- Gupta, B. B., Perez, G. M., Agrawal, D. P., & Gupta, D. (2019). *Handbook of Computer Networks and Cyber Security: Principles and Paradigms*. Springer Nature.
- Harkness, B., & Black, M. (2020, June 28). How to Dispute Errors and Mistakes in Your Credit Reports. Retrieved August 28, 2020, from Credit Card Insider website: <https://www.creditcardinsider.com/learn/errors-mistakes-in-your-credit-report/>
- Harrison, O. (2018, October 15). What is Survey Data Processing? Retrieved December 13, 2020, from Displayr website: <https://www.displayr.com/what-is-survey-data-processing/>
- Hasan, R., & Yurcik, W. (2006). A Statistical Analysis of Disclosed Storage Security Breaches. *Proceedings of the Second ACM Workshop on Storage Security and Survivability*, 1–8. New York, NY, USA: ACM. <https://doi.org/10.1145/1179559.1179561>

- Hayes, A. (2020, March 3). Reading Into Stratified Random Sampling. Retrieved March 4, 2021, from Investopedia website:
https://www.investopedia.com/terms/stratified_random_sampling.asp
- HDS. (2019, February). *Hitachi Virtual Storage Platform (VSP) Encryption Engine Non-Proprietary CryptFolIgPrSa 1p4h0ic- 2M odule Security Policy*. HDS. Retrieved from <https://csrc.nist.gov/csrc/media/projects/cryptographic-module-validation-program/documents/security-policies/140sp2462.pdf>
- Heder, B. (2014, November 13). Redundancy and failover and HA, oh my! Retrieved December 6, 2020, from Network World website:
<https://www.networkworld.com/article/2847353/redundancy-and-failover-and-ha-oh-my.html>
- Henry, G. T. (1998). *Practical Sampling*. Sage.
- Herrera, T. (2019, November 24). You're Tracked Everywhere You Go Online. Use This Guide to Fight Back. (Published 2019). *The New York Times*. Retrieved from <https://www.nytimes.com/2019/11/24/smarter-living/privacy-online-how-to-stop-advertiser-tracking-opt-out.html>
- HHS. (2009, September 10). The Security Rule [Text]. Retrieved December 7, 2020, from HHS.gov website: <https://www.hhs.gov/hipaa/for-professionals/security/index.html>
- Hibbard, E. (2016). *Intro to Encryption and Key Management: Why, What and Where?* SNIA.
- Hibbard, E. A. (2011). *SNIA Storage Security Best Practices*. SNIA.
- Hibbard, E. A. (2014). *Best Practices for Cloud Security and Privacy*. SBIA. Retrieved from http://www.snia.org/sites/default/orig/DSI2014/presentations/Security/Hibbard_Best-Practices-for-Cloud-Security-and-Privacy%20final.pdf
- Hibbard, E. A. (2015). *SNIA Storage Security Best Practices*. SNIA.
- Hibbard, E., & Rivera, T. (2014, September). *Reforming EU Data Protections...No Ordinary Sequel*. SNIA. Retrieved from https://www.snia.org/sites/default/files/Hibbard_EU-Data-Protection%20_v2_Final_0.pdf

- Hirsch, D. D. (2013). The Law and Policy of Online Privacy: Regulation, Self-Regulation, or Co-Regulation? *Seattle University Law Review*, 34, 42.
- Hou, H., Yu, J., & Hao, R. (2019, Nay). *Cloud storage auditing with deduplication supporting different security levels according to data popularity*. ScienceDirect. Retrieved from <https://www.ezproxy.dsu.edu:2065/science/article/pii/S1084804519300669>
- Hoven, J. van den, Blaauw, M., Pieters, W., & Warnier, M. (2019). Privacy and Information Technology. In *The Stanford Encyclopedia of Philosophy* (Summer 2020). Metaphysics Research Lab, Stanford University. Retrieved from <https://plato.stanford.edu/archives/sum2020/entries/it-privacy/>
- Hubbert, S. (2011). *Data center storage; cost-effective strategies, implementation, and management*. SNIA.
- IBM. (2019a, September). *IBM Storage Insights: Security Guide*. IBM. Retrieved from https://www.ibm.com/support/knowledgecenter/SSQRB8/com.ibm.spectrum.si.doc/IBM_Storage_Insights_Security_Guide.pdf
- IBM. (2019b, October 14). File-storage. Retrieved December 13, 2019, from <https://www.ibm.com/cloud/learn/file-storage>
- IDG. (2017). Data & Analytics Survey, IDG. *Hubspot*, 7.
- Imam, F. (2019, May 22). What Is the CIA Triad and Why Is It Important for Cybersecurity? Retrieved August 11, 2020, from Logsign website: <https://blog.logsign.com/what-is-the-cia-triad-and-why-is-it-important-for-cybersecurity/>
- Intellectus Statistics. (2020). *Intellectus Statistics* [Online computer software]. Retrieved from <https://analyze.intellectusstatistics.com/>
- ISO. (2015). *ISO/IEC 27040:2015 Information technology—Security techniques—Storage security*. ISO. Retrieved from <http://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/04/44/44404.html>
- Isreal, G. D. (2005). Determining Sample Size. *University of Florida IFAS Extension*. Retrieved from <https://www.tarleton.edu/academicassessment/documents/samplesize.pdf>

- Jian-hua, Z., & Nan, Z. (2011, August). *Cloud Computing-based Data Storage and Disaster Recovery*. IEEE. Retrieved from <https://ieeexplore.ieee.org/abstract/document/6041774>
- Jovanovic, V., & Mirzoev, T. (2010). Teaching Storage Infrastructure Management and Security. *2010 Information Security Curriculum Development Conference*, 41–44. New York, NY, USA: ACM. <https://doi.org/10.1145/1940941.1940952>
- Kahan, B. C., Rehal, S., & Cro, S. (2015). Risk of selection bias in randomised trials. *Trials*, *16*(1), 405. <https://doi.org/10.1186/s13063-015-0920-x>
- Kent, K., & Souppaya, M. P. (2006). *Guide to computer security log management* (No. NIST SP 800-92; 0 ed., p. NIST SP 800-92). Gaithersburg, MD: National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-92>
- Klebnikov, S. (2019, November 6). Companies With Security Fails Don't See Their Stocks Drop As Much, According To Report. Retrieved December 24, 2020, from Forbes website: <https://www.forbes.com/sites/sergeiklebnikov/2019/11/06/companies-with-security-fails-dont-see-their-stocks-drop-as-much-according-to-report/>
- Komprise. (2009, September 19). What is Unstructured Data? Data Management Glossary. Retrieved March 4, 2021, from Komprise website: https://www.komprise.com/glossary_terms/unstructured-data/
- Krahn, R., Trach, B., Vahldiek-Oberwagner, A., Knauth, T., Bhatotia, P., & Fetzer, C. (2018). Pesos: Policy Enhanced Secure Object Store. *Proceedings of the Thirteenth EuroSys Conference on - EuroSys '18*, 1–17. Porto, Portugal: ACM Press. <https://doi.org/10.1145/3190508.3190518>
- Krzyzanowski, P. (2009). Cryptographic communication and authentication. *Rutgers University*, 25.
- Kumar, D. S., Rawat, U. S., Jasra, S. K., & Jain, A. K. (2009). Efficient methodology for implementation of Encrypted File System in User Space. *ArXiv:0908.0551 [Cs]*. Retrieved from <https://arxiv.org/pdf/0908.0551.pdf>
- Kwon, J., & Johnson, M. E. (2018, December). *Meaningful Healthcare Security: Does "Meaningful-Use" Attestation Improve Information Security Performance?* EBSCOhost.

Retrieved from <https://www.econinfosec.org/archive/weis2014/papers/KwonJohnson-WEIS2014.pdf>

Li, L., Qian, K., Chen, Q., Hasan, R., & Shao, G. (2016). Developing Hands-on Labware for Emerging Database Security. *Proceedings of the 17th Annual Conference on Information Technology Education*, 60–64. New York, NY, USA: ACM.
<https://doi.org/10.1145/2978192.2978225>

Li, W., Yang, Y., & Yuan, D. (2015). Reliability Assurance of Big Data in the Cloud. In *Reliability Assurance of Big Data in the Cloud* (pp. 9–17). Boston: Morgan Kaufmann.
<https://doi.org/10.1016/B978-0-12-802572-7.00002-6>

LII Cornell Law School. (2012, April 11). 18 U.S. Code § 1350—Failure of corporate officers to certify financial reports. Retrieved December 7, 2020, from LII / Legal Information Institute website: <https://www.law.cornell.edu/uscode/text/18/1350>

Liu, S. (2017, April). Global data storage problems 2016-2017. Retrieved December 13, 2019, from Statista website: <https://www.statista.com/statistics/752840/worldwide-data-storage-problems/>

Local Government Association. (2020, June 18). General Data Protection Regulation (GDPR) | Local Government Association. Retrieved December 7, 2020, from <https://www.local.gov.uk/our-support/guidance-and-resources/general-data-protection-regulation-gdpr>

LTCC. (2009, October 17). Small Table of z-values for Confidence Intervals. Retrieved December 10, 2020, from <http://www.ltconline.net/greenl/courses/201/estimation/smallConfLevelTable.htm>

Mackey, D. (2008, February). Basel II's impact on information security. Retrieved December 6, 2020, from SearchSecurity website: <https://searchsecurity.techtarget.com/tip/Basel-II-s-impact-on-information-security>

Madison, T. (2020, July 1). Data is Now the World's Most Valuable Commodity—Here's How You Can Own Yours. Retrieved January 17, 2021, from Medium website: <https://medium.com/decentranet/data-is-the-most-valuable-commodity-on-earth-heres-how-you-can-own-your-data-94acb7a9ee75>

- Marek.Z. (2013, December 10). RPO, RTO, WRT, MTD...WTH?! Retrieved November 30, 2020, from Default Reasoning website: <https://defaultreasoning.com/2013/12/10/rpo-rto-wrt-mtdwth/>
- Market Research Guy. (2020, October 5). Types of data measurement scales: Nominal, ordinal, interval, and ratio. Retrieved December 16, 2020, from My Market Research Methods website: <https://www.mymarketresearchmethods.com/types-of-data-nominal-ordinal-interval-ratio/>
- McGavisk, T. (n.d.). The Positive and Negative Implications of GDPR. Retrieved July 14, 2020, from Tbs website: <https://www.timedatasecurity.com/blogs/the-positive-and-negative-implications-of-gdpr>
- McKay, K. A., Polk, W. T., & Chokhani, S. (2014, April). *Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations*. NIST. Retrieved from <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-52r1.pdf>
- McMinn, M. (2009a, June). *Information Management – Extensible Access Method (XAM) – Part 2: C API*. SNIA. Retrieved from https://www.snia.org/sites/default/files/XAM_C_API_v1.01.pdf
- McMinn, M. (2009b, June). *Information Management – Extensible Access Method (XAM) – Part 3: Java API*. SNIA. Retrieved from https://www.snia.org/sites/default/files/XAM_Java_API_v1.01.pdf
- McMinn, M. (2009c, June). *Information Management—Extensible Access Method (XAM)—Part 1: Architecture*. SNIA. Retrieved from https://www.snia.org/sites/default/files/XAM_Arch_v1.01.pdf
- Menard, S. (2009). *Logistic regression: From introductory to advanced concepts and applications*. Thousand Oaks, CA: Sage Publications.
- Meslhy, E., Abd elkader, H., & Eletriby, S. (2013). Data Security Model for Cloud Computing. *Journal of Communication and Computer* 10 (2013) 1047-1062, 10, 1047–1062. <https://doi.org/10.13140/2.1.2064.4489>

- Mulligan, S. P., Freeman, W. C., & Linebaugh, C. D. (2019, March). *Data Protection Law: An Overview*. Congressional Research Service. Retrieved from <https://fas.org/sgp/crs/misc/R45631.pdf>
- Naeem, T. (2020, April 30). Data Integrity in a Database—Why Is It Important. Retrieved August 28, 2020, from Astera website: <https://www.astera.com/type/blog/data-integrity-in-a-database/>
- National Advertising Initiative. (2018, May). Enforcement | NAI: Network Advertising Initiative. Retrieved November 11, 2020, from ENFORCEMENT website: <https://www.networkadvertising.org/code-enforcement/enforcement>
- NCSL. (2019, May 29). Data Security Laws | Private Sector. Retrieved January 17, 2021, from Data Security Laws | Private Sector website: <https://www.ncsl.org/research/telecommunications-and-information-technology/data-security-laws.aspx>
- Nelson, A. (2017, October 4). How to Secure Private Data Stored and Accessed in the Cloud. Retrieved August 19, 2020, from Principles for Digital Development website: <https://digitalprinciples.org/resource/howto-secure-private-data-cloud/>
- Neuhäuser, M., Lehmann, N., Nonnemacher, M., & Stausberg, J. (2006). An attempt at data verification in the EACTS Congenital Database: Data before and after verification differ significantly. *European Journal of Cardio-Thoracic Surgery*, *30*(4), 691–691. <https://doi.org/10.1016/j.ejcts.2006.06.023>
- NGINX. (2018, September 27). What Is Load Balancing? How Load Balancers Work [Industry]. Retrieved August 19, 2020, from NGINX website: <https://www.nginx.com/resources/glossary/load-balancing/>
- NIST. (2006). *Minimum security requirements for federal information and information systems* (No. NIST FIPS 200; p. NIST FIPS 200). Gaithersburg, MD: National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.FIPS.200>
- Noordzij, M., Dekker, F. W., Zoccali, C., & Jager, K. J. (2011). Sample Size Calculations. *Nephron Clinical Practice*, *118*(4), c319–c323. <https://doi.org/10.1159/000322830>

- NortonLifeLock. (n.d.). Help secure your accounts with these strong password tips [Industry]. Retrieved August 17, 2020, from Norton website: <https://us.norton.com/internetsecurity-how-to-how-to-secure-your-passwords.html>
- OCR. (2015, September 10). HIPAA for Professionals [Text]. Retrieved December 1, 2020, from HHS.gov website: <https://www.hhs.gov/hipaa/for-professionals/index.html>
- OECD. (2011). *The Evolving Privacy Landscape: 30 Years After the OECD Privacy Guidelines*. <https://doi.org/10.1787/5kgf09z90c31-en>
- Office of Attorney General California. (2018, October 15). California Consumer Privacy Act (CCPA). Retrieved December 8, 2020, from State of California—Department of Justice—Office of the Attorney General website: <https://oag.ca.gov/privacy/ccpa>
- Paik, J.-Y., Choi, J.-H., Jin, R., Wang, J., & Cho, E.-S. (2018, October 8). *A Storage-level Detection Mechanism against Crypto-Ransomware*. 2258–2260. ACM. <https://doi.org/10.1145/3243734.3278491>
- Park, S.-W., Lim, J., & Kim, J. N. (2015). A Secure Storage System for Sensitive Data Protection Based on Mobile Virtualization. *International Journal of Distributed Sensor Networks*, 11(2), 929380. <https://doi.org/10.1155/2015/929380>
- Patterson, C. (2018, November). *Why your current disaster recovery strategy may not cover compliance*. Navisite. Retrieved from <https://drj.com/wp-content/uploads/2019/08/WP-Why-your-DR-strategy-may-not-cover-compliance-1-1-18.pdf>
- PCI SSC. (2010, October). *PCI DSS Quick Reference Guide Understanding the Payment Card Industry Data Security Standard version 3.2*. PCI SSC. Retrieved from https://www.pcisecuritystandards.org/documents/PCI_DSS-QRG-v3_2_1.pdf?agreement=true&time=1570331249834
- PCI SSC. (2015, June). *Payment Card Industry (PCI) Data Security Standard*. PCI SSC. Retrieved from <https://www.pcisecuritystandards.org/>
- PCI SSC. (2017, May). *PCI Data Security Standard (PCI DSS)*. PCI SSC. Retrieved from <https://www.pcisecuritystandards.org>

- PCI SSC. (2018a, June). *Payment Card Industry (PCI) Data Security Standard Report on Compliance. PCI DSS v3.2 Template for Report on Compliance. Revision PDF Free Download*. PCI. Retrieved from <https://docplayer.net/29382158-Payment-card-industry-pci-data-security-standard-report-on-compliance-pci-dss-v3-2-template-for-report-on-compliance-revision-1-0.html>
- PCI SSC. (2018b, June). *The Prioritized Approach to Pursue PCI DSS Compliance*. PCI SSC. Retrieved from https://www.pcisecuritystandards.org/documents/Prioritized-Approach-for-PCI-DSS-v3_2_1.pdf
- Peter P. Swire & DeBrae Kennedy-Mayo. (2018). *U.S. Private-Sector Privacy: Law and Practice for Information Privacy Professionals* (2nd ed.). International Association of Privacy Professionals.
- Peters, J. (2020, January 20). Data Privacy Guide: Definitions, Explanations and Legislation | Varonis. Retrieved November 9, 2020, from Inside Out Security website: <https://www.varonis.com/blog/data-privacy/>
- Pituch, K. A., & Stevens, J. P. (2015). *Applied Multivariate Statistics for the Social Sciences* (6th ed.). New York: Routledge. Retrieved from <https://doi.org/10.4324/9781315814919>
- Ponemon Institute, & IBM Security. (2020, September 8). Cost of a Data Breach Report 2020. Retrieved December 24, 2020, from <https://www.ibm.com/security/digital-assets/cost-data-breach-report/#/pdf>
- Poojary, P. (2019, March 12). Understanding Object Storage and Block Storage use cases | Cloud Academy Blog. Retrieved December 13, 2019, from Cloud Academy website: <https://cloudacademy.com/blog/object-storage-block-storage/>
- Porter, Y., Piscopo, T., & Marke, D. (2014, August 14). Object Storage versus Block Storage: Understanding the Technology Differences. Retrieved December 13, 2019, from Druva website: <https://www.druva.com/blog/object-storage-versus-block-storage-understanding-technology-differences/>
- Pottier, R., & Menaud, J.-M. (2017, April 24). *Privacy-aware Data Storage in Cloud Computing*. 405–412. Scitepress. Retrieved from <https://www.scitepress.org/Link.aspx?doi=10.5220/0006294204050412>

- PricewaterhouseCoopers. (2018, November 6). PwC's Global Data and Analytics Survey 2016: Big Decisions. Retrieved January 17, 2021, from PwC website:
<https://www.pwc.com/us/en/services/consulting/analytics/big-decision-survey.html>
- Razali, N. M., & Wah, Y. B. (2011). Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, (2(1)), 21–33.
- Rice, M. E., & Harris, G. T. (2005). *Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r*. American Psychological Association. Retrieved from
<https://doi.org/10.1007/s10979-005-6832-7>
- Riley, P. (2019, June). Good Data Analysis | ML Universal Guides. Retrieved December 13, 2020, from Google Developers website: <https://developers.google.com/machine-learning/guides/good-data-analysis>
- Roell, K. (2019, July 22). What Is Grading on a Curve? Retrieved December 16, 2020, from ThoughtCo website: <https://www.thoughtco.com/grading-on-a-curve-3212063>
- Rouse, M. (2019, June). What is blockchain storage? Retrieved December 15, 2019, from SearchStorage website: <https://searchstorage.techtarget.com/definition/blockchain-storage>
- Rouse, M., Cook, R., & Wigmore, I. (2012, July). What is Multipath I/O (MPIO)? - Definition from WhatIs.com. Retrieved August 19, 2020, from WhatIs.com website:
<https://whatis.techtarget.com/definition/Multipath-I-O-MPIO>
- Rouse, M., Sullivan, E., Posey, B., Diamantis, C., & Yamamura, Y. (2020, February). What is RAID (Redundant Array of Independent Disks)? Retrieved August 19, 2020, from SearchStorage website: <https://searchstorage.techtarget.com/definition/RAID>
- RSI Security. (2018, December 20). Penalties for Non-Compliance with FISMA (and how to avoid them). Retrieved December 6, 2020, from RSI Security website:
<https://blog.rsisecurity.com/penalties-for-non-compliance-with-fisma-and-how-to-avoid-them/>

- Saltis, S. (2020, November 5). GDPR Fines: Everything You Need To Know. Retrieved November 11, 2020, from Core dna website: <https://www.coredna.com/blogs/gdpr-fines>
- Sarkar, M. K., & Chatterjee, T. (2014). *Enhancing Data Storage Security in Cloud Computing Through Steganography*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.428.6647&rep=rep1&type=pdf>
- Schaffer, K. (2019). ITL BULLETIN MAY 2019 FIPS 140-3 Adopts ISO/IEC Standards. *NIST*, 3.
- Schneider, J. W. (2009). PREVENTING DATA BREACHES: ALTERNATIVE APPROACHES TO DETER NEGLIGENT HANDLING OF CONSUMER DATA. *Boston University School of Law*, 15, 25.
- Schopmeyer, A., & Somasundaram, G. (2009). *Information Storage and Management: Storing, Managing, and Protecting Digital Information*. O’O’Reilly`.
- Schopmeyer, K. (2017). *Automation of SMI-S managed storage systems with Pywbem*. 47.
- Schulz, G. (2011). *Cloud and Virtual Data Storage Networking* (1st ed.). CRC Press.
- Siponen, M., Puhakainen, P., & Vance, A. (2019). Can individuals• neutralization techniques be overcome? A field experiment on password policy | Elsevier Enhanced Reader. *Sciense Direct*, 88. <https://doi.org/10.1016/j.cose.2019.101617>
- SIRE. (2019, May 14). GDPR one year on: The positive and negative implications. Retrieved November 11, 2020, from SIRE website: <https://www.sire.co.uk/gdpr-one-year-on-the-positive-and-negative-implications/>
- SNIA. (2008, June). *ISCSI Management API*. SBIA. Retrieved from https://www.snia.org/sites/default/files/iSCSIManagementAPI_v2.0.pdf
- SNIA. (2009, March). *Common RAID Disk Data Format Specification*. SBIA. Retrieved from https://www.snia.org/sites/default/files/SNIA_DDF_Technical_Position_v2.0.pdf
- SNIA. (2010a, March). *Multipath Management API*. SNIA. Retrieved from https://www.snia.org/sites/default/files/MMA_Technical_Position_v1.1.pdf

- SNIA. (2010b, June). *Hypervisor Storage Interfaces for Storage Optimization White Paper*.
SNIA. Retrieved from https://www.snia.org/sites/default/files/HSI_Copy_Offload_WP-r12.pdf
- SNIA. (2012, March). *Architectural Model for Data Integrity*. SNIA. Retrieved from
http://snia.org/sites/default/files/Data_Integrity_Architectural_Model_v1.0.pdf
- SNIA. (2014, November). *TLS Specification for Storage Systems*. SNIA. Retrieved from
<http://www.snia.org/sites/default/files/TLSSpec-v1%20Technical%20Position.pdf>
- SNIA. (2015a, March). *Cloud Data Management Interface (CDMITM) Version 1.1.1*. SNIA.
Retrieved from https://www.snia.org/sites/default/files/CDMI_Spec_v1.1.1.pdf
- SNIA. (2015b, March). *Sanitization*. SNIA. Retrieved from
<https://www.snia.org/sites/default/files/SNIA-Sanitization-TechWhitepaper.pdf>
- SNIA. (2015c, August). *Storage Security: Encryption and Key Management*. SNIA. Retrieved
from https://www.snia.org/sites/default/files/technical_work/SecurityTWG/SNIA-Encryption-KM-TechWhitepaper.R1.pdf
- SNIA. (2016a). *Storage Security: Fibre Channel Security*. SNIA. Retrieved from
https://www.snia.org/sites/default/files/technical_work/SecurityTWG/SNIA-FC-Security-TechWhitepaper.160906.pdf
- SNIA. (2016b, August). *Storage Security: An overview as applied to storage management
Version 1*. SNIA. Retrieved from
https://www.snia.org/sites/default/files/technical_work/SecurityTWG/SNIA-Storage-Mgmt-Security-TechWhitepaper.pdf
- SNIA. (2016c, December). *Self-contained Information Retention Format (SIRF) Specification*.
SNIA. Retrieved from
https://www.snia.org/sites/default/files/technical_work/SIRF/SIRF_V1_Technical_Position.pdf
- SNIA. (2017a, June). *NVM Programming Model (NPM)*. SNIA.

- SNIA. (2017b, October). *Data Protection Best Practices*. SNIA. Retrieved from https://www.snia.org/sites/default/files/DPCO/Data%20Protection%20BP%20White%20Paper%20Final%20v1_0.pdf
- SNIA. (2018a, March). *Storage Networking Industry Association*. SNIA. Retrieved from <https://www.snia.org/sites/default/files/security/SNIA-Data-Protection-TechWhitepaper.pdf>
- SNIA. (2018b, November). *Contact us via LiveChat!* SNIA. Retrieved from <https://secure.livechatinc.com/>
- SNIA. (2019, May). *Linear Tape File System (LTFS) Format Specification*. SNIA. Retrieved from https://www.snia.org/sites/default/files/technical_work/LTFS/LTFS_Format_v2.5_Technical_Position.pdf
- Solove, D. (2015, November 13). The Growing Problems with the Sectoral Approach to Privacy Law. Retrieved November 11, 2020, from TeachPrivacy website: <https://teachprivacy.com/problems-sectoral-approach-privacy-law/>
- Stanganelli, J. (2019, November 20). Compliance and Data Privacy Regs that Affect IT Security [Security]. Retrieved April 16, 2020, from ESecurity Planet website: <https://www.esecurityplanet.com/network-security/security-compliance.html?b>
- Stobierski, T. (2019, August 26). The Advantages of Data-Driven Decision-Making | HBS Online. Retrieved January 16, 2021, from Business Insights—Blog website: <https://online.hbs.edu/blog/post/data-driven-decision-making>
- Subha, T., & Jayashri, S. (2017, January). *Efficient privacy preserving integrity checking model for cloud data storage security*. IEEE. Retrieved from <https://ieeexplore.ieee.org/document/7951745>
- Swire, P., & Kennedy-Mayo, D. (2018). *u.s. Private-sector privacy law and practice for information privacy professionals* (2nd ed.). International Association of Privacy Professionals.

- Talib, A. M., Atan, R., Murad, M. A. A., & Abdullah, R. (2010). A FRAMEWORK OF MULTI AGENT SYSTEM TO FACILITATE SECURITY OF CLOUD DATA STORAGE. *International Conference on Cloud Computing & Virtualization*, 241–258.
- Tang, Y., Li, Q., Li, D., Li, Z., Zhang, M., Jee, K., ... Xu, F. (2018). NodeMerge: Template Based Efficient Data Reduction For Big-Data Causality Analysis. *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security - CCS '18*, 1324–1337. Toronto, Canada: ACM Press. <https://doi.org/10.1145/3243734.3243763>
- Techopedia. (2017, June 2). What is a Data Center? - Definition from Techopedia. Retrieved December 13, 2019, from Techopedia.com website:
<https://www.techopedia.com/definition/349/data-center>
- Temple. (2016, July 28). Policy: Procedure for Revocation of System Access Upon Termination | Temple ITS. Retrieved August 19, 2020, from Temple University website:
<https://its.temple.edu/procedure-revocation-system-access-upon-termination>
- Thomson Reuters. (2020, September 18). Understanding Data Privacy – A Compliance Strategy Can Mitigate Cyber Threats. Retrieved December 2, 2020, from Understanding data privacy: A compliance strategy can mitigate cyber threats website:
<https://legal.thomsonreuters.com/en/insights/articles/understanding-data-privacy-a-compliance-strategy-can-mitigate-cyber-threats>
- Tozzi, C. (2020, April 5). Data Availability—6 Reasons Your Data May Become Unavailable. Retrieved November 30, 2020, from Precisely website:
<https://www.precisely.com/blog/data-availability/reasons-data-availability-unavailable>
- TrueVault. (2019, January 19). How much do HIPAA violations cost? - TrueVault. Retrieved December 7, 2020, from <https://www.truevault.com/resources/compliance/how-much-do-hipaa-violations-cost>
- Tunggal, A. T. (2020, August 5). What is SOX Compliance? Overview, Requirements, and Controls | UpGuard. Retrieved December 6, 2020, from <https://www.upguard.com/blog/sox-compliance>

- University of Baltimore. (2004, September 12). Questionnaire Design and Surveys Sampling [Academic]. Retrieved January 16, 2021, from Questionnaire Design and Surveys Sampling website: <http://home.ubalt.edu/ntsbarsh/Business-stat/stat-data/Surveys.htm>
- US Bureau of Labor Statistics. (2020). QCEW Second Quarter 2020 Response Rates. Retrieved December 10, 2020, from <https://www.bls.gov/cew/response-rates/cew-response-rates-second-quarter-2020-2019.htm>
- US DOE. (2018, February 7). FERPA | Protecting Student Privacy. Retrieved December 7, 2020, from FERPA website: <https://studentprivacy.ed.gov/node/548/>
- Vandersreen, J. (2019). The Disadvantages of Encrypted Files. Retrieved August 22, 2020, from It Still Works website: <https://itstillworks.com/disadvantages-encrypted-files-2597.html>
- Vanhove, J. (2015, November 2). Causes and consequences of unequal sample sizes. Retrieved December 10, 2020, from <https://janhove.github.io/design/2015/11/02/unequal-sample-sized>
- Vasilopoulos, D., Elkhyaoui, K., Molva, R., & Onen, M. (2018). POROS: Proof of Data Reliability for Outsourced Storage. *Proceedings of the 6th International Workshop on Security in Cloud Computing*, 27–37. New York, NY, USA: ACM.
<https://doi.org/10.1145/3201595.3201600>
- Veleva, P. (2019). *PERSONAL DATA SECURITY FOR SMART SYSTEMS AND DEVICES WITH REMOTE ACCESS*. EBSCOhost. Retrieved from <http://www.uni-sz.bg/tsj/Volume%2017,%202019,%20Supplement%201,%20Series%20Social%20Sciences/4/za%20pe4at/144.pdf>
- Verizon. (2020). Verizon: Data Breach Investigations Report 2020. *Computer Fraud & Security*, 2020(6), 4. [https://doi.org/10.1016/S1361-3723\(20\)30059-2](https://doi.org/10.1016/S1361-3723(20)30059-2)
- Walkowski, D. (2019, July 9). What Is The CIA Triad? Retrieved August 19, 2020, from F5 Labs website: <https://www.f5.com/labs/articles/education/what-is-the-cia-triad>
- Walsh, K. (2018, November 8). Checklist For FedRAMP Requirements. Retrieved December 6, 2020, from Reciprocity website: <https://reciprocitylabs.com/checklist-for-fedramp-requirements/>

- Wang, C., Wang, Q., Ren, K., & Lou, W. (2010). Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing. *2010 Proceedings IEEE INFOCOM*, 1–9. <https://doi.org/10.1109/INFCOM.2010.5462173>
- "Wang, H., Yang, D., Duan, N., Guo, Y., & Zhang, L. (2019, March). *Medusa: Blockchain Powered Log Storage System*. IEEE. Retrieved from <https://www.ezproxy.dsu.edu:2119/stamp/stamp.jsp?tp=&arnumber=8663935>
- Wang, X., & Cheng, G. (2018, May). *Design and Implementation of Universal City Disaster Recovery Platform*. IEEE. Retrieved from <https://ieeexplore.ieee.org/document/8469356>
- Weins, K. (2018, January 17). Compare Top Public Cloud Providers: AWS vs Azure vs Google. Retrieved December 13, 2019, from Flexera Blog website: <https://www.flexera.com/blog/cloud/2018/01/compare-top-public-cloud-providers-aws-vs-azure-vs-google/>
- Westfall, P. H., & Henning, K. S. S. (2013). *Understanding advanced statistical methods* (1st ed.). Taylor & Francis. Retrieved from 13: 978-1466512108
- Wikipedia. (2019). Unstructured data. In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Unstructured_data&oldid=929003912
- Willett, M. (2012). *Implementing Stored-Data Encryption*. 50.
- Woodard, R. L. (2004). Is Your Medical Information Safe? A Comparison of Comprehensive and Sectoral Privacy and Security Laws. *Indiana International & Comparative Law Review*, 15(1), 147–182. <https://doi.org/10.18060/17834>
- Xu, Y. (2018, December). *Section-Blockchain: A Storage Reduced Blockchain Protocol, the Foundation of an Autotrophic Decentralized Storage Architecture*. IEEE. Retrieved from <https://ieeexplore.ieee.org/document/8595065>
- Yeager, K. (2014, May 18). LibGuides: SPSS Tutorials: Independent Samples t Test. Retrieved January 5, 2021, from <https://libguides.library.kent.edu/SPSS/IndependentTTest>
- Zamboni, J. (2018, May 15). The Advantages of a Large Sample Size. Retrieved December 15, 2020, from Sciencing website: <https://sciencing.com/advantages-large-sample-size-7210190.html>

- Zheng, Q., Li, Y., Chen, P., & Dong, X. (2018, December). *An Innovative IPFS-Based Storage Model for Blockchain*. IEEE. Retrieved from <https://ieeexplore.ieee.org/document/8609675>
- Zhou, J. (2014). On the Security of Cloud Data Storage and Sharing. *Proceedings of the 2Nd International Workshop on Security in Cloud Computing*, 1–2. New York, NY, USA: ACM. <https://doi.org/10.1145/2600075.2600087>
- Zhou, L., Varadharajan, V., & Gopinath, K. (2016). A Secure Role-Based Cloud Storage System For Encrypted Patient-Centric Health Records. *The Computer Journal*, 59(11), 1593–1611. <https://doi.org/10.1093/comjnl/bxw019>
- Zhu, Y., Wang, H., Hu, Z., Ahn, G., Hu, H., & Yau, S. S. (Dec2010). Dynamic Audit Services for Integrity Verification of Outsourced Storage. In *Clouds,*” in *Proc. ACM Symposium on Applied Computing (SAC), 2011*, 1550–1557. ACM. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.700.8372&rep=rep1&type=pdf>
- Zyskind, G., Nathan, O., & Pentland, A. (2015, Jul). *Decentralizing Privacy: Using Blockchain to Protect Personal Data*. IEEE. Retrieved from <https://www.ezproxy.dsu.edu:2119/stamp/stamp.jsp?tp=&arnumber=7163223>

APPENDIX A: REGULATIONS STUDIED

BASEL II

Official website: <https://www.bis.org/publ/bcbsca.htm>

Year: 2004

Data Protection Model: Sectoral Financial

Industry: Financial Banking

Governing Body: Basel Committee on Bank Supervision

Data Type: Financial, PII

Description: Set of international banking regulations with uniform guidelines defined by the Basel Committee on Bank Supervision (CFI, 2020)

Requirements:

- Capital adequacy: requires banks to maintain a minimum capital adequacy requirement of 0.8% of their risk-weighted assets (RWA) (BIS, 2004)
- Supervisory review: This pillar enables various regulatory bodies to deal with risks like systemic risk, liquidity, and legal risks (BIS, 2004)
- Market discipline: requires banks to disclose information regarding risk exposures and capital adequacy (BIS, 2004)

Non-Compliance: Suspension of operational licenses, fines, and restrictions of payment to directors and shareholders (CFI, 2020)

Security Controls: (Mackey, 2008)

- Requires backup or DR: N/A
- Encryption: N/A
- Opt-in or Opt-out: N/A
- Notification in case of incidents: Yes, required to disclose risk exposures
- System and Data Protection: Yes
- Risk assessment, monitoring, and mitigation: Yes
- Regular Auditing: Yes

Principles:

1. Authentication: No
2. Authorization: No
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
7. Accessibility: No

California Consumer Privacy Act (CCPA)

Official website: <https://www.oag.ca.gov/privacy/ccpa>

Year: 2020

Data Protection Model: Comprehensive California residents

Governing Body: State of California - Department of Justice - Office of the Attorney General

Industry: All

Data Type: Personal Data

Description: Gives consumers in California more control over the information that businesses collect about them (Office of Attorney General California, 2018)

Requirements: (Office of Attorney General California, 2018)

- Mandate's businesses disclose to any collection, use, or sharing of personal data
- Mandates a simple method for individual opt-out of the sale of personal data
- Must inform the consumer of the right to delete their data and delete upon request
- Delete all personal information at the request of the consumer and must notify the consumer of the right to that request
- The act also prohibits discrimination of any user exercising their CCPA rights

Non-Compliance: (Becerra, 2018)

- Legal action by the California Attorney General or civil class action suits to pay statutory damages of up to \$750 per California resident and incident
- Recommends civil penalties of up to \$7500 per intentional violation while unintentional violations carry a penalty of \$2500 per record

Security Controls: (Office of Attorney General California, 2018)

- Requires backup or DR: N/A
- Encryption: Yes
- Opt-in or Opt-out: Opt-out
- Notification in case of incidents: Yes
- Vulnerability Assessment and Remediation: Yes
- Establish and Monitor Audit Logs: Yes
- Antivirus and Malware Defense: Yes
- Access controls: Yes, Authorization and Authentication

Principles:

1. Authentication: Yes
2. Authorization: Yes
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: No
7. Accessibility: Yes

Family Educational Rights and Privacy Act (FERPA)

Official website: <https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html>

Year: 1974

Data Protection Model: Sectoral Education

Governing Body: US Department of Education

Industry: Education

Data Type: Students' and Minors' data, PHI *may also include as part of student records (ASTHO, 2014)

Description: Protects the privacy of student education records. Applies to all educational institutions that receive funds from the US Department of Education.

Requirements: (US DOE, 2018)

- School network must be protected
- Must honor 'Do not share' request
- Access to parents for students under 18 years of age
- Reliability: Records will be updated if errors are discovered

Non-Compliance: Loss of federal funding or other administrative actions. It is important to note that FERPA only applies to schools that receive funding from the US Department of Education. (US DOE, 2018)

Security Controls: (US DOE, 2018)

- It does not require specific technical security controls. However, it does make recommendations. This link is a comprehensive data security checklist posted by DOE:
 - https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Data%20Security%20Checklist_0.pdf
- Notification in case of incidents: Yes
- Policy Access Controls: Yes, based on the age of consent
- Data and System protection: Yes
- Opt-in or Opt-out: N/A
- Physical security: ensuring unauthorized users cannot access computing resources
- Inventory keeping: Yes
- Authentication: Yes
- Audit and compliance monitoring: Yes

Principles:

1. Authentication: Yes
2. Authorization: Yes
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
7. Accessibility: No

Federal Financial Institutions Examination Council (FFIEC)

Official website: <https://www.ffiec.gov/>

Year: 1979

Data Protection Model: Sectoral Financial

Governing Body: Federal Financial Institutions Examination Council (FFIEC)

Industry: Banking and Finance

Data Type: PII and Financial Data

Description: An interagency body consisting of several banking regulators mandated to establish principles and standards that promote financial institutions' consistency. Banking regulators including; FRB, FDIC, NCUA, OCC, and CFPB (FFIEC, 1996)

Requirements: (FFIEC, 1996)

- Provide clear and conspicuous notices on privacy policies and practices to all customers
- Categorization of non-public personal information collected
- Notify the customer of their opt-out right to any sharing of their information to third party entities
- Ensure strict data, systems, and communication protection

Non-Compliance: (FFIEC, 1996)

- Subpoena or summons from the Federal or state authorities
- Civil, criminal, or regulatory investigations
- Fines and Penalties depending on the agency that regulates the institution and the severity

Security Controls: (FFIEC, 1996)

- Requires backup or DR: Yes
- Encryption: Yes, of online transaction processing (OLTP)
- Opt-in or Opt-out: Opt-out
- Notification in case of incidents: Yes
- Authorization and Authentication: Yes, Multifactor authentication
- Cybersecurity risk assessment and monitoring: Yes
- Penetration Testing and Vulnerability scanning: Yes
- System patch management: Yes
- Information Audits: Yes

Principles:

1. Authentication: Yes
2. Authorization: Yes
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
7. Accessibility: Yes

Federal Information Security Management Act (FISMA)

Official website: <https://www.cisa.gov/federal-information-security-modernization-act>

Year: 2002

Data Protection Model: Sectoral Government

Governing Body: Office of Management and Budget (OMB) (CISA, 2019)

Industry: Federal Government

Data Type: Covered Government Information (CGI) and all sensitive data

Description: Law requiring all federal agencies to develop, document, and implement an agency-wide program related to all information systems (RSI Security, 2018)

Requirements: (NIST, 2006)

- Requires complete inventory of systems and assets
- Required to classify all systems based on data stored and accessed
- Requires strict security controls
- Required to conduct three-tier risk assessments; organizational, business process, and information system level
- Requires FISA certification

Non-Compliance: Loss of federal funding, Potential government hearings and scrutiny, censure from future contracts, and lack of customer trust, which will negatively affect the business (Bansal, 2020)

Security Controls: (NIST, 2006)

- Requires backup or DR: Yes
- Encryption: Yes
- Opt-in or Opt-out: N/A
- Notification in case of incidents: Yes
- Access controls and logs: Yes
- Firewalls and Antivirus: Yes
- Risk assessment, vulnerability scans, and penetration tests: Yes
- Media Protection: Yes
- Contingency planning: Yes
- Awareness and Training: Yes
- Personnel security: Revoke user's access when terminated or transferred
- Audit and accountability: Yes

Principles:

1. Authentication: Yes
2. Authorization: Yes
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
7. Accessibility: Yes

Federal Risk and Authorization Management Program (FedRAMP)

Official website: <https://www.fedramp.gov/>

Year: 2011

Data Protection Model: Sectoral Government

Governing Body: Office of Management and Budget (OMB)

Industry: Federal Government

Data Type: Covered Government Information (CGI) and Controlled Unclassified Information (CUI) for Cloud storage

Description: provides a standardized approach to security assessment authorization and continuous monitoring for cloud products and services within the federal Government (Walsh, 2018). Federal agencies, Cloud Services Providers, and third-party assessment organizations. (Walsh, 2018)

Requirements:

- Cloud services providers are required to implement security controls following FIPS 199 (Walsh, 2018)
- Continuous Monitoring (ConMon) program to for consistent vulnerability scans (Walsh, 2018)
- Develop a roadmap to meet the controls which may require architectural changes (Dunham, 2018)
- Document them under FedRAMP System Security Plan (SSP) (Dunham, 2018)

Non-Compliance: Loss of certification, suspension, or revocation to do business with government agencies (FedRAMP PMO, 2018)

Security Controls: (Dunham, 2018)

- Requires backup or DR: Yes
- Encryption: Yes
- Opt-in or Opt-out: N/A
- Notification in case of incidents: Yes
- Firewalls: Yes
- Risk assessment and monitoring: Yes
- Authentication and Authorization: Yes
- Vulnerability assessment and penetration testing: Yes
- Systems, data, and communications protection: Yes, attained by abiding by policies developed by the NIST
- System Categorization: Yes, as either low or moderate impact levels

Principles:

1. Authentication: Yes
2. Authorization: Yes
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
7. Accessibility: Yes

General Data Protection Regulation (GDPR)

Official website: <https://gdpr.eu>

Year: 2018

Data Protection Model: Comprehensive EU residents

Governing Body: Information Commissioner's Office (ICO) (Local Government Association, 2020)

Industry: All

Data Type: Personal Data

Description: Sets out a legal framework on data protection principles, rights, and obligations of businesses worldwide on behalf of EU members (Frankenfield, 2020)

Requirements: (Saltis, 2020)

- Legal basis for all collection and processing of EU citizens' personal data
- Mandates that EU members give disclosure of data breaches promptly

Non-Compliance: Judicial action and monetary penalties of up to 2% of global turnover or 10 million Euros, whichever is higher (GDPR.EU, 2018a)

Security Controls: (GDPR.EU, 2018a)

- Requires backup or DR: N/A
- Encryption: Yes
- Opt-in or Opt-out: Opt-in
- Notification in case of incidents: Yes
- Auditing: Yes

Principles:

1. Authentication: No
2. Authorization: No
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: No
7. Accessibility: No

Gramm–Leach–Bliley Act (GLBA)

Official website: <https://www.ftc.gov/tips-advice/business-center/privacy-and-security/gramm-leach-bliley-act>

Year: 1999

Data Protection Model: Sectoral Financial

Governing Body: Federal Trade Commission (FTC)

Industry: Financial

Data Type: PII and Financial Data

Description: requires financial institutions that offer consumers financial services like loans, investment consultancy, and insurance to disclose their information sharing practices and safeguard sensitive data (FTC, 2002)

Requirements: (FTC, 2002)

- Issue customers and consumers clear note of privacy policies
- Right to opt-out of sharing data with third parties
- Require clear privacy and security documented plans

Non-Compliance: (FTC, 2002)

- Financial institutions face fines of \$100,000 per violation
- Individuals in charge of non-compliant institutions and found in violation face fines of \$10,000 for each violation or a jail term of up to 5 years

Security Controls:

- Requires backup or DR: N/A
- Encryption: Yes
- Opt-in or Opt-out: Opt-out
- Notification in case of incidents: Yes
- Risk assessment, Monitoring, and Testing: Yes
- Systems and Data Protection: Yes
- Access Controls: Yes, the system should ensure accountability of users and keep track of the information they access
- Authorization and Authentication: Yes

Principles:

1. Authentication: Yes
2. Authorization: Yes
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
7. Accessibility: No

Health Insurance Portability and Accountability Act (HIPAA)

Official website: <https://www.hhs.gov/hipaa/index.html>

Year: 1996

Data Protection Model: Sectoral Healthcare

Governing Body: S. Department of Health and Human Services (HHS)

Industry: Healthcare

Data Type: PHI

Description: outlines a national legal standard to data protection and security of PHI by covered entities (CDC, 2019)

Requirements: (CDC, 2019)

- Requires covered entities to ensure confidentiality, Integrity and Availability of PHI
- Safeguard it against anticipated threats like unauthorized access, use, and disclosures
- Certify compliance of their workforce

Non-Compliance: Recommends monetary penalties of up to \$50000 per violation. Furthermore, violators face class actions that could result in jail time (TrueVault, 2019)

Security Controls: (HHS, 2009)

- Requires backup or DR: Yes
- Encryption: Yes
- Opt-in or Opt-out: N/A *providers not accepting insurance or government funds can opt-out
- Notification in case of incidents: Yes

Principles:

1. Authentication: No
2. Authorization: No
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
7. Accessibility: No

Payment Card Industry Data Security Standard (PCI DSS)

Official website: <https://www.pcisecuritystandards.org/>

Year: 2006

Data Protection Model: Co-Regulatory and Self-Regulatory

Governing Body: PCI Security Standards Council (PCI SSC)

Industry: Financial Credit Cards

Data Type: Financial Data, PII

Description: Defines standards to ensure a secure environment is maintained by all organizations that process, store or transmit credit card information (PCI SSC, 2018b)

Non-Compliance: (PCI SSC, 2017)

- Fines and penalties: up to \$100,000 monthly or legal action that can be in the millions
- Loss of credit card services

Requirements: (PCI SSC, 2017)

- Organizations are required to protect all cardholder data
- Use and maintain network security
- Encrypt all transmitted data
- Restrict physical and data access
- Maintaining access logs
- Authorization information should be isolated from application data

Security Controls: (PCI SSC, 2018a)

- Requires backup or DR: N/A
- Encryption: Yes
- Opt-in or Opt-out: N/A
- Notification in case of incidents: Yes
- Access controls and logs: Yes
- Firewalls: Yes
- Antivirus: Yes
- Risk assessment, vulnerability scans, and penetration tests: Yes

Principles:

1. Authentication: No
2. Authorization: Yes
3. Privacy: Yes
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
7. Accessibility: Yes

Sarbanes–Oxley Act (SOX)

Official website: <https://sarbanes-oxley-act.com/>

Year: 2002

Data Protection Model: Sectoral Financial

Industry: Financial Investment

Governing Body: United States Securities and Exchange Commission (SEC)

Data Type: Financial, PII

Description: Established to protect investors from fraudulent financial reporting by corporations (Tunggal, 2020)

Requirements: (Tunggal, 2020)

- Management assessment, establish internal controls and reporting mechanisms
- Requires companies to keep adequate record-keeping for at least 5 years. Including electronic communications
- Companies need to undergo accuracy and control audits
- Requires senior company executives to personally sign off and certify that financial records are compliant with SEC regulations to assert responsibility in case of fraud.
- Companies should issue reports to investors in real-time without omissions or untrue statements
- Companies should protect any presumed whistleblowers within the company
- Criminal penalties for falsifying or altering financial reports

Non-Compliance: (LII Cornell Law School, 2012)

- Fines, prison time, or both, can be up to five million dollars and 20 years' incarceration
- Removal from the public stock exchange
- Voiding of directors and senior executives' insurance policies

Security Controls:

- Requires backup or DR: Yes, for a certain period of years.
- Encryption: N/A
- Opt-in or Opt-out: N/A
- Notification in case of incidents: N/A
- Compliance Auditing: Yes
- Access controls: Yes, physical access and role-based access control for electronic data
- Data Protection: Yes
- Data Classification: Yes, depending on data sensitivity and the applicable regulations

Principles:

1. Authentication: No
2. Authorization: Yes
3. Privacy: No
4. Reliability: Yes
5. Verification: Yes
6. Recoverability: Yes
1. Accessibility: Yes

APPENDIX B: SURVEY

Data Privacy and Security Survey Consent for Anonymous Data Privacy and Security Survey

You are invited to participate in a research study. In order to participate, you must be 18 years old. Taking part in this research project is voluntary. Please take time to read this entire message and ask questions before deciding whether to take part in this survey.

The goal of this research study is to identify common issues, threats, risks, and trends in stored data security and privacy, as well as assess their impact on the activities of private, financial, education, healthcare, government, and other organizations.

This goal can be, in part, achieved by collecting feedback from industry professionals to examine trends, concerns, issues, and opinions as it relates to stored data security. The study will correlate the responses to related data compliance regulations to determine the gaps or shortcomings to stored data with a goal to help organizations identify areas for further development and improvement of data privacy and security.

All responses to this survey are strictly confidential, anonymous, and stored in a secure, offline location. No specific, personally identifiable information will be asked. All requested information has been reviewed and approved by Dakota State University's (DSU) Institutional Review Board (IRB).

Upon completion of the review of all questionnaires received, the data will be compiled and analyzed and used for data privacy and security research in my work as a cybersecurity graduate student at DSU. No information that personally identifies you or your organization will be used in any analysis or reported in any form.

There is no reward or other compensation for participating in this research, and no penalty for withdrawing after participation has begun. Information collected from participants who withdraw from the survey will not be used in analyses or retained.

1.

***I agree to participate in the research study and confirm that I am at least 18 years old. I understand the purpose and nature of this study and I am participating voluntarily. I understand that I can withdraw from the study at any time, without any penalty or consequences.**

Agree Disagree

2. What region are you in? (Select 1)

- USA
- Europe
- Asia-Pacific
- Canada
- Latin America
- I prefer not to answer
- Other (please specify) _____

3. What is your organization's industry or business type? (Select 1)

- Finance & Financial Services
- Advertising & Marketing
- Business Support & Logistics
- Airlines & Aerospace (including Defense)
- Health Care & Pharmaceuticals
- Education
- Manufacturing
- Government
- Retail & Consumer Durables
- Telecommunications, Technology, Internet & Electronics
- Nonprofit
- Utilities, Energy, and Extraction
- I prefer not to answer
- Other (please specify) _____

4. How many employees are there in your organization? (Select 1)

- 50 or less employees
- 51 to 249 employees
- 250 to 999 employees
- 1000 to 10,000 employees
- Over 10,000 employees
- I don't know
- I prefer not to answer
- Other (please specify) _____

5. What type of information/data does your organization collect, use, and store? (Select all that apply)

- Personally Identifiable Information (PII) – i.e., Individuals driver's license, government id number, address, etc.
- Protected Health Information (PHI) – i.e., patients medical/health records, medications, treatments, etc.
- Personal data - i.e., age, gender, likes/dislikes, sexual orientation, religion, family, online social platforms info, diet, political views, pets, etc.
- Employee information
- Customer information or data
- Financial data – i.e., credit card data, bank accounts, etc.
- Student information
- Data for minors, i.e., children under the age of 18 years
- Your organization's intellectual property - i.e., trade secrets, procedures, designs, developed code, etc.
- None
- I don't know
- I prefer not to answer
- Other (please specify) _____

6. Which laws and regulations do you consider relevant to your organization? (select all that apply)

- General Data Protection Regulation (GDPR)
- California Consumer Privacy Act (CCPA)
- Health Insurance Portability and Accountability Act (HIPAA)
- Payment Card Industry Data Security Standard (PCI DSS)
- Federal Information Security Management Act (FISMA)
- Family Educational Rights and Privacy Act (FERPA)
- Sarbanes–Oxley Act (SOX)
- BASEL II
- Federal Risk and Authorization Management Program (FedRAMP)
- Gramm–Leach–Bliley Act (GLBA)
- Federal Financial Institutions Examination Council (FFIEC)

- None
 I don't know
 I prefer not to answer
 Other (please specify) _____

7. How significant are the following consequences of unlawful, unauthorized, or accidental types of data incidents to your organization?

(On the scale of 1 to 5, where 1 is "Not significant", 5 is "Very significant")

Reputation and brand damage – bad or embarrassing press

1	2	3	4	5	N/A
---	---	---	---	---	-----

Reduced revenue or customer loss

1	2	3	4	5	N/A
---	---	---	---	---	-----

Loss of trust on the part of interested parties

1	2	3	4	5	N/A
---	---	---	---	---	-----

Litigation / legal proceedings

1	2	3	4	5	N/A
---	---	---	---	---	-----

Deterioration of relations with employees'

1	2	3	4	5	N/A
---	---	---	---	---	-----

Regulatory actions/sanctions or fines

1	2	3	4	5	N/A
---	---	---	---	---	-----

Loss of competitive advantage (for example, due to loss of intellectual property)

1	2	3	4	5	N/A
---	---	---	---	---	-----

8. How important is improving data security in ensuring the following activities?

(On the scale of 1 to 5, where 1 is "Not Important", 5 is "Very Important")

Reputation and Brand Protection

1	2	3	4	5	N/A
---	---	---	---	---	-----

Intellectual Property Protection

1	2	3	4	5	N/A
---	---	---	---	---	-----

Personal data protection

1	2	3	4	5	N/A
---	---	---	---	---	-----

Providing support when launching a new service

1	2	3	4	5	N/A
---	---	---	---	---	-----

Regulatory Compliance

	1	2	3	4	5	N/A
Compliance with internal policies						
	1	2	3	4	5	N/A
Improving IT management and operations						
	1	2	3	4	5	N/A
Increased stakeholder confidence						
	1	2	3	4	5	N/A
Interaction with external suppliers						
	1	2	3	4	5	N/A
Learning New and Emerging Technologies						
	1	2	3	4	5	N/A
Assistance in mergers, acquisitions, and sales						
	1	2	3	4	5	N/A

9. Which area of your organization is your primary function or role? (Select 1)

- Sales, Marketing or Business Development
 Analyst
 Project Management
 Strategy/Planning
 Accounting/Auditing
 Administrative
 Customer Service
 Engineering
 Human Resources
 Health Care Provider
 Finance
 Educator (e.g., teacher, lecturer, professor, trainer)
 Information Technology
 Research and Development
 management
 Consulting
 Quality Assurance
 research
 Student
 Other (please specify) _____

10. What are the requirements to access your organization's systems? (Select all that apply)

- Username/password
 Biometric identification (fingerprint, facial recognition, other)
 Two-factor authentication
 I don't know
 I prefer not to answer

Other (please specify) _____

11. Regarding system security management, which options are true: (Select all that apply)

- I am required to change my password regularly
- My organization requires complex passwords (i.e., 10+ characters, must contain upper and lower alpha chars, numbers, and special characters)
- My organization will not allow me to use previous passwords
- My organization has an online self-service website where I can update my password
- My organization has dedicated helpdesk support if I have computer or access issues
- My organization never requires me to change my password
- I don't know
- I prefer not to answer

12. What impact has regulatory requirements (HIPAA, PCI, GDPR, CCPA, etc.) had on the effectiveness of data security in your organization? (Select 1)

- Significant increase in the effectiveness of ensuring information security as a result of compliance with regulatory requirements
- Moderate increase in efficiency
- No change
- Information Security Efficiency decreased
- I don't know
- I prefer not to answer
- Other (please specify) _____

13. The following general organizational measures are in place to protect data: (Select all that apply)

- The person responsible for the organization's data protection has been appointed
- Audits are done
- A policy has been developed and published to protect information (personal data)
- Developed regulations, orders, or instructions are in place
- Developed terms of reference for the creation of an information protection system are in place
- Developed technical design of information security system in place
- Information security tools have been introduced
- My organization's computer (i.e., work laptop) is updated and secured regularly
- I don't know
- I prefer not to answer
- Other (please specify) _____

14. If I am working on a required system at my organization and find something that isn't accurate (i.e., typo or some other incorrect information):

- I can update or modify without any issue
- I need to open a support ticket or notify someone else
- There isn't anything I can do
- I don't know
- I prefer not to answer
- Other (please specify) _____

15. If a person outside my organization (member, student, customer, patient, etc.) wants or needs to update, modify, or delete personal data or information: (Select all that apply)

- They can update or modify without issue
- They can delete without issue
- They must contact my organization
- It cannot be updated, modified, or deleted

- My organization doesn't store outside individual's information
- I don't know
- I prefer not to answer
- Other (please specify) _____

16. Who in your organization oversees information security issues? (Select 1)

- CEO, CISO, CIO, CTO or DPO
- Other Security / Privacy officer
- Compliance Director / Director of Personal Data
- Protection Technical Director
- Head of Internal Audit
- Risk management director
- IT Specialist
- Information Security Specialist
- Network / System Administrator
- I don't know
- I prefer not to answer
- Other (please specify) _____

17. What if you delete an email and then you need to restore it:

- I can easily do this
- I can contact support, and they can restore it
- It's gone, and I can't restore it
- I don't know if it can be restored
- I prefer not to answer

18. If I delete or lose a file (I.e., document, spreadsheet, etc.):

- I can easily do this
- I can contact support, and they can restore it
- It's gone, and I can't restore it
- I don't know if it can be restored
- I prefer not to answer

19. If a disaster or incident is affecting my "normal" tasks, I am trained by my organization, and I know what I need to do and whom I need to contact:

- True
- False
- I prefer not to answer

20. Which of the following things are accessible to you? (Select all that apply)

- I must be in my office to access my organization's applications
- I can access my organization email from my personal computer
- I can access organization email from my smartphone or tablet
- I have an organization laptop and remote access
- I have VPN access to my organization's systems and applications
- I can access my organization's applications with any browser over the Internet
- I don't know
- I prefer not to answer

21. On a scale of 1 to 5, where 1 - "Strongly Disagree", 5 - "Strongly Agree", rate the following statements:

My organization makes data security and privacy the highest priority

1	2	3	4	5	N/A
---	---	---	---	---	-----

My organization, makes changing privacy and security settings easy

1	2	3	4	5	N/A
---	---	---	---	---	-----

APPENDIX C: IRB APPROVALS

To: Pam Rowland, Howard Goodman

Date: 07/02/2020

Project Title: Data Privacy Compliance A false sense of Data Storage Security

Approval #: 20200702-1

I have reviewed your application and related documents and determined that your project falls outside of definitions used in federal regulations that govern the protections of human subjects in research under 45 CFR 46.102(e) and (l). Your proposed conduct of a survey in a manner that does not seek any personally identifiable information led to this decision.

While your project has been determined exempt from IRB review, you will need to conduct all project activities in accordance with DSU, State of South Dakota, and federal rules and policies.

You must immediately contact the IRB if:

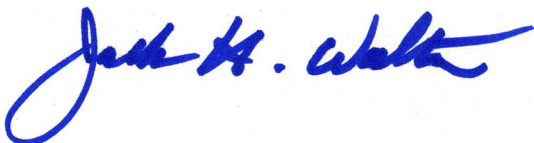
- the nature of your project changes in any way that relates to human subjects
- any event occurs that could indicate the release of participant identities, no matter how small
- any other unexpected or adverse event occurs in relation to your project
- any material changes are made to your research plan or process

IRB rules require that you provide an update on the status of your project (1) when the research is complete (a notice of closure), or (2) by one year from yesterday (07/01/2020), whichever comes first.

If you have questions about this determination or the conditions stated in this memo, please contact us at 605-256-5100 or irb@dsu.edu.

Best wishes for success in your research and related endeavors.

Yours truly,



Jack H. Walters, Ph.D.
Professor of Management and
Chair, Institutional Review Board

APPENDIX D: INDUSTRY CATEGORIES

The below list is all unique responses for question 3: *“What is your organization’s industry or business type?”* All responses are shown as selected or typed in by participants with typos and spelling errors.

#	Q3. Industry Response	Industry Category	Q3. Industry Response	Industry Category
1	No Response	Disqualified	civil engineer	Business Services
3	Engineering	Technology	legal/life coach	Business Services
5	Government	Government	N'gineer	Technology
7	Consumer Retail	Retail	Structural Engineering - Building Construction	Business Services
9	Entertainment	Other	Environmental Consulting	Other
11	Retirred	Disqualified	Motorcycle Service and Repair shop	Retail
13	PNTA	Other	career planning	Business Services
15	Education	Education	Auto interiors	Retail
17	Transportation	Other	Consulting and Training	Education
19	Finance & Financial Services (retired)	Financial	Maintenance and construction	Other
21	Health care	Healthcare	Investigation Services	Business Services
23	Retired	Disqualified	Land Surveying and Engineering	Other
25	Student of Health Information Technology/Administration	Technology	Design - Build	Business Services
27	Cyber Security	Technology	Residential construction (home building)	Other
29	Construction	Other	Laundromat	Retail
31	Manufacturing	Other	Outdoor Recreation	Other
33	Nonprofit	Nonprofit	Standby power	Utility
35	retired - consumer product safety	Retail	hospitality/travel	Other
37	ENTERTAINMENT	Other	Electrician	Other
39	Telco, ISP or Tech	Technology	Home Improvement	Other
41	Health & Safety	Healthcare	General Construction	Other
43	Pest control	Other	Wholesale automotive parts distribution	Retail
45	Mental Health	Healthcare	Demolition/Construction	Other
47	Church/Ministry	Nonprofit	Vehicle Rental	Retail
49	real estate	other	Arts	Other
51	Media (newspaper)	Other	Construction Equipment Sales	Retail
53	Financial	Financial	Shipyards	Other
55	Hospitality	Business Services	retied	Disqualified
57	author	Other	auto tech	Other
59	Retired supt of schools	Education	Lawn care	Other
61	Horticultural communications	Other	Rsidential Remodeling	Other
63	Agriculture	Other	Patents	Business Services
65	Real Estate Brokerage	Other	Business services	Business Services
67	Transportation	Other	Electrical Contracting	Other
69	Music Production	Other	Food Service	Other
71	Utilities	Utility	Coin Dealer	Other
73	Airlines	Other	architect and builder	Business Services

75	Marketing	Business Services	Tile and Grout cleaning	Other
77	Hospitality/Lodging	Business Services	interior design	Retail
79	Business Support	Business Services	Service - fireplaces	Retail
81	Own various businesses	Other	Service industry-Hairsylist	Retail
83	Retired from Financial Services	Financial	Commuter Railroad	Government
85	Law Firm	Business Services	Own a construction company	Other
87	Dretired	Disqualified	Software development & consulting across many business types	Technology
89	Consulting	Business Services	Environmental Consulting and training	Education
91	Hospitality Management	Business Services	automated transit systems-people movers	Other
93	news media	Other	Religious Organization	Nonprofit
95	Farming	Other	Contractor Restoration	Other
97	Law	Business Services	hair salon/beauty	Retail
99	I am retired from City gov't work, i.e: Water/Wastewater Treatment	Utility	General contracting	Business Services
101	Food service	Other	Furniture	Retail
103	Architecture & Interior Design	Business Services	Facility Planning, Design and Project Management	Business Services
105	constructions	Other	consulting- advanced materials	Business Services
107	I don't have an organization.	Disqualified	Residential and commercial building inspections	Government
109	Homemaker	Disqualified	technical services	Technology
111	Information Security Audits	Technology	Consumer Consumables (wine)	Retail
113	I don't have an organization	Disqualified	Cattleman	Other
115	Retired	Disqualified	Landscaping Services	Other
117	Wellness and health education	Education	Automotive Parts and Repair	Retail
119	Health&Fitness	Healthcare	Home Repair	Other
121	Entertainment/Performance/Music	Other	Landscape Construction and property maintenance services	Other
123	Woodland fire	Other	Massage Therapy	Other
125	NONE	Disqualified	fire district	Other
127	I am retired, no business.	Disqualified	Landscaping	Other
129	Quilting thread and notions	Disqualified	Repair Shop	retail
131	Pet Care	Other	Art studio	Other
133	Privacy, information security & risk management	Technology	venture capital	Financial
135	Disabled, but prior to that County law enforcement	Government	wholesale trade	Retail
137	Retired Healthcare professional	Healthcare	Meeting design and facilitation	Other
139	Brewing	Other	Landscape Installation	Other
141	Environmental	Other	import export	Business Services
143	Prefer not to answer	Disqualified	Sculpture casting	Other
145	Metals distribution/	Other	Fitness & wellness	Other
147	Medium, Death Duola, Reverend, NLP practitioner, Care giver	Disqualified	Food & Beverage	Retail
149	I have been retired for 25 years as a Lab Tech.	Disqualified	Upholstery and furniture repair	Retail
151	Master Automotive Mechanic	Other	Real estate & development	Other

153	Engineering and Construction	Business Services	Lodging	Retail
155	Engineering	Technology	radon testing and mitigation and home inspections	Government
157	Retired from retail management	Retail	Construction trade	Other
159	Real estate and legal services	Business Services	Was Retail, now Scientific Research	Other
161	household	Disqualified	Designer/fabricator	Other
163	retired clinical psychologist, private practice	Healthcare	RACQUET AND FITNESS	Other
165	MINISTRY	Nonprofit	Rental Equipment	Retail
167	Semi retired	Disqualified	Legal services	Business Services
169	Private Sector Consultant in the field of Urban Planning	Nonprofit	Restoration of fire and water damage	Other
171	Insurance	Business Services	Staffing	Business Services
173	Software	Technology	Auto repair	Other
175	Entertainment Industry	Other	personal email linked to an outdated web page, but people still send me questions concerning its content.	Disqualified
177	Public service	Government	Real Estate - Home Inspection	Business Services
179	Medical Devices - Retired	Healthcare	Wholesale Distribution	Retail
181	I'm retired from Government	Government	Software & legal consulting	Technology
183	Professional services	Business Services	Computer Technician	Technology
185	Insurance Claims	Business Services	survey research	Education
187	Retired for about 22 years, Howard	Disqualified	Professional Engineering Consulting	Technology
189	Retired college prof	Education	Lawn and Landscape Maintenance	Other
191	Retired but maintaining four professional certifications in cybersecurity.	Technology	Management consulting, mainly for tech startups	Business Services
193	None? I deliver pizzas	Disqualified	Religious	Nonprofit
195	Real Estate	Other	Executive Search supporting nonprofit executive recruitment nationwide since 2001.	Nonprofit
197	Photography	Other	Church	Nonprofit
199	Legal	Business Services	Uniform and Linen rental	Other
201	home services	Other	Real estate sales	Retail
203	Management consulting	Business Services	Real estate investments	Financial
205	writer of novels	Other	Locksmith	Other
207	Law enforcement	Government	Travel	Other
209	Retired Government	Government	Ecological services	Other
211	Community organizations	Nonprofit	bookbinding as a service business	Retail
213	Restaurant	Other	Commercial transactional printing	Retail
215	Automotive	Other	catholic parish, non profit	Nonprofit
217	Mental health outpatient/ private practice	Healthcare	Presbyterian Church	Nonprofit
219	spiritual preception	Other	Book Publishing	Other
221	Construction	Other	Home remodeling	Retail
223	Rental House Landlord	Other	Religious institution	Nonprofit
225	tourism	Other	Service	Other

227	Ordained religious leader	Nonprofit	Lawn maintenance	Other
229	Search & Rescue	Government	architectural services	Business Services
231	Automotive Services	Other	Tourism & Hospitality	Other
233	Retired School Transportation Director	Education	Architecture, professional services	Business Services
235	Consulting, professional	Business Services	Data Broker	Business Services
237	Insurance	Business Services	Arts Representation	Other
239	retired military	Government	environmental chemistry	Other
241	Energy, Agriculture, and Timber	Utility	Engineering consulting	Technology
243	not in any business (retired)	Disqualified	Brewery, restaurant, Online retail sales,	Retail
245	Education, Design, and Technology	Education	Photography	Other
247	Retired Military Retired Handicaped work supervisor	Government	Home Building	Other
249	I'm a student	Education	e-commerce	Retail
251	Food Container Mfg	Other	design	Other
253	Publishing	Other	Training Boiler Operators	Education
255	meat processor	Other	musician's union	Other
257	ranching	Other	Tree Farm and Nursery	Other
259	self employed writer of poems, quotes	Other	Religious/church	Nonprofit
261	I'm retired	Disqualified	Real Estate Development	Business Services
263	Media	Other	Business Management Consultant- All industries	Business Services
265	Retired Banker	Financial	Chemistry Testing Laboratory	Other
267	Landlord & Tenant .	Other	Christian/ Baptist Church	Nonprofit
269	Maintenance	Other	Analyti SW developer	Technology
271	Retired Government	Government	software	Technology
273	Retired/Hairdresser/Multi-Media AS degree	Disqualified	Childcare	Other
275	Church	Nonprofit	Pet Srevicees	Other
277	Fabric and quilting	Other	Nonprofit consulting	Nonprofit
279	Lobster fisherman	Other	Graphic Design / Fine Art Photography	Other
281	retired finance	Financial	heavy equipment repair	Other
283	I am retired	Disqualified	Travel and Tourism	Business Services
285	Architectural & Design Services	Business Services	Medical School	Education
287	Commercial fishing	Other	Church, religious organization	Nonprofit
289	Beauty Industry	Other	Food Industry	Other
291	disabled	Disqualified	Arts marketing	Other
293	I am retired	Disqualified	Physical Therapy Recruiter	Business Services
295	Homemaker	Disqualified	Professional Staffing	Business Services
297	currently unemployed, previously with Higher Ed & public health	Education	lawyer	Business Services
299	Student	Education	Ag Sales	Retail
301	Barber	Other	Human Resources Consultancy	Business Services
303	Entertainment (Radio program host)	Other	Real estate appraisal	Other
305	Law/lawyer/legal services	Business Services	Mobile cranes sales and rentals	Retail
307	Investment group	Financial	Design Consulting	Business Services
309	Resale/recycling computers and electronics	Technology	Church, religion, theology	Nonprofit

311	Horse Ranching	Other	Service	Business Services
313	Real estate/construction	Other	Research	Education
315	Heating & Cooling Repair & Installation	Other	Executive Recruiting	Business Services
317	CPA Firm	Business Services	Computer Hardware / Software / Support for Temporary Personnel	Technology
319	architect	Business Services	Travel and Leisure - Consumer Discretionary	Retail
321	Public Relations	Other	Professional Employer Organization (PEO)	Business Services
323	financial planning software	Technology	Real estate and mortgage	Business Services
325	Graphic Design	business Services	Art	Other
327	Air Conditioning Contractor Sales / Service	Retail	Human services	Business Services
329	HVAC	Other	Fitness	Other
331	Distribution	Business Services	I am a real-estate broker.	Other
333	HVAC sales, service and installation	Retail	Pet grooming	Other
335	repair and installation of plumbing and heating equipment	Other	Attorney	Business Services
337	HVAC service	Other	Private Sporting & Social Club	Other
339	Management	Business Services	Construction/ steel	Other
341	3D Artist	Other	travel (tour operator)	Other
343	Insurance consulting, rm	Business Services	Fitness	Other
345	engineering and design	Technology	Commercial Printing	Technology
347	Services, law	Business Services	Portrait photography	Other
349	legal services (law firm)	Business Services	Legal Services/Law Office	Business Services
351	Law Firm - Legal Services	Business Services	Engineering Consultant	Technology
353	Residential Design	Business Services	Information Technology	Technology
355	Legal services	Business Services	General Contracting (Home Repairs & Remodeling)	Other
357	Technology, Multimedia, Broadcasting	Technology	Agriculture/Food (Wineries and Vineyards)	Other
359	Professional Services - Architectural	Business Services	Other	Other
361	lawyer - retired but only recently	Business Services	Automotive - Engineering, Manufacturing & Motorsports	Other
363	Lega/Attorney	Business Services	Professional Services (ie. Sales and Marketing)	Business Services
365	law office	Business Services	LABOR ORGANIZATION	Nonprofit
367	Architecture/Engineering/Construction	Business Services	industrial, commercial sales and service	Retail
369	Crafts	Other	Construction and Mechanical Services	Other
371	SaaS	Technology	IT Managed Services	Technology
373	Consultant	Business Services	Distribution Vending Machines. Founder of mfg from South Dakota	Retail
375	computer I.T.	Technology	Real Estate Dev & Mgmt	Other
377	Full Service Restaurant	Other	sales and engineering	Technology
379	personal services/pet care	Other	Fire and emergency services	Other
381	automotive repair	Other	Funeral supply company	Retail
383	architecture	Business Services	Construction Products	Retail
385	Public Transportation	Government	Research - computer science	Education

387	I'm retired and have a small business selling products online and in person.	Retail	entertainment/journalism	Other
389	Service - Automotive Repair	Other	Attorney in private practice	Business Services
391	retired chiropractic	Healthcare	image licensing	Other
393	RV park - travel industry	Other	Photography, custom framing, graphic & web design	Other
395	I am a trial attorney	Business Services	Trade Association	Other
397	law (retired)	Business Services	real estate agent	Other
399	Realtor	Other	Media, specifically photojournalist	Other
401	Retired software engineer	Technology	Engineering Services	Technology
403	Computer Repair	Technology	Design Build Remodeling	Other
405	Publisher	Other	auto repair and sales	Retail
407	Hospitality	Other	Wholesale Nursery	Retail
409	Restaurant	Other	on-site furniture repair	Retail
411	Maritime (Boat retail and repair)	Retail	Indian Tribe	Nonprofit
413	Business & IT Consulting	Technology	Residential Plumbing Service	Retail
415	RARE BOOKSELLER	Retail	Plumbing	Other
417	Strategic Communications Consulting	Technology	Real Estate- Appraisals	Other
419	Wholesale	Retail	Mechanical Engineering Consulting	Technology
421	Chemical distribution	Retail	Science & Research	Education
423	Consultant, Biz Services, Tax Preparation, misc	Business Services		

APPENDIX E: ORGANIZATION SIZE

The below list is all individual responses for question 4: *“How many employees are there in your organization?”*
All responses are shown as selected or typed in by participants with typos and spelling errors.

#	Q4. Number Employees	Organization Size	Q4. Number Employees	Organization Size
1	251 to 999	Medium	Retired was 2000 district employees	Large
3	No Response	Disqualified	Me and my wife.	Small
5	51 to 250	Medium	not in any business (retired)	Other
7	I'm retired and don't know how I was selected for this survey	Disqualified	see previous	Disqualified
9	1000 to 9999	Large	milions	Large
11	10K+	Large	I'm retired	Other
13	1 to 50	Small	disabled	Disqualified
15	PNTA	Other	4	Small
17	Unknown	Other	tetired	Disqualified
19	Retired	Other	3	Small
21	1-Me	Other	Retired	Disqualified
23	Retitrd	Other	I was a solo practitioner	Other
25	1 - just me	Other	One	Other
27	1	Other	sole practitioner	Other
29	I do not work	Other	I have independent contractors no employees	Small
31	Disabled	Other	No employees; just me and my wife.	Other
33	0	Other	two	Small
35	unemployed	Disqualified	1 part time	Small
37	self-employed	Other	Just me	Other
39	Refer to Question #3.	Other	Solo operation	Other
41	I work for the state.	Large	2	Small
43	N/A	Disqualified	no employees sole proprietor	Other
45	No employees	Other	Self	Other
47	I am unemployed. I have no organization	Disqualified	8	Small
49	NA	Other	1 person	Other
51	self-employed - just me	Other	sole-proprietier	Other
53	I am retired	Other	just me & my partner, and consultants as needed.	Small
55	One - Individual Private Practice	Other	no employees, all volunteers	Other
57	no longer working	Other	We don't have employees per se per IRS definitions.	Other
59	none	Other	5 or less	Small
61	Just me. I consult for an investment bank.	Other	No employees, this is my own personal email.	Disqualified
63	just one im a business owner/operator	Other	60 volunteers; no paid staff	Small
65	Its not my organization. I work in a hospital 35 employees on inpatient floor	Small	We are small but subsidiary part of a worldwide organization. Church entity.	Small
67	self employed	Small	Retired DC. At max, I had 4 assistants	Small

69	retired, formerly worked in small orgs w/50 people or less	Small	8 Volunteer Directors	Small
71	I am self-employed and I have no staff	Other	1 full-time, 1 part-time	Small
73	I am now retired	Other	5	Small
75	locally or internationally?	Other	Myself	Other
77	250,000	Large	One full-time, one part-time	Small
79	Presently unemployed. Last employer 51 - 249.	Medium	None. All volunteer.	Other
81	Being retired, none	Other	small firm with 4 people	Small
83	Other	Disqualified	i'm getting annoyed so i'm leaving	Disqualified
85	1, myself	Other	no employees; all-volunteer nonprofit	Small

APPENDIX F: ROLE CATEGORY

The below list is all individual responses for question 9: “Which area of your organization is your primary function or role?” All responses are shown as selected or typed in by participants with typos and spelling errors.

#	Q9. Working Role	Role Category	Q9. Working Role	Role Category
1	IT	Technical	Owner, jack of all trades Pretty much everything. I'm a sole practitioner with one employee (a secretary)	Leadership
3	No Response	Disqualified		Leadership
5	Student	Non-Technical	Partner	Leadership
7	reporter	Non-Technical	Litigation	Non-Technical
9	Analyst	Technical	Professional Services	Technical
11	Educator	Educator	Law Firm Managing Partner	Leadership
13	Management	Leadership	lawyer owner	Leadership
15	Accounting/Auditing	Non-Technical	Legal representation of clients Attorney & Chief Privacy Officer	Non-Technical
17	Sales or Marketing	Non-Technical	legal services	Leadership
19	PM	Leadership	Lawyer	Non-Technical
21	Health Care	Non-Technical	owner, sole decision maker	Non-Technical
23	author	Non-Technical	legal	Leadership
25	Administrative	Non-Technical	Owner - sole proprietor	Non-Technical
27	Consultant	Technical	Prosecuting Attorney	Leadership
29	HR	Non-Technical	Attorney/Owner	Non-Technical
31	CS	Technical	All operations As a small business owner, all of the above.	Leadership
33	Song Writer/Producer	Non-Technical	owner/all of the above	Non-Technical
35	Finance	Non-Technical	N'gineering	Leadership
37	Engineering	Technical	Boss, I do everything listed.	Technical
39	Corrections facility	Leadership	Business Owner/Innkeeper Everything. I'm a sole proprietor.	Leadership
41	Owner/manager	Leadership	Everything Emergency 911	Leadership
43	Owner	Leadership	Telecommunications Owner/operator which covers most of the above	Leadership
45	R&D	Technical	As sole owner, management and IT and customer service, etc principle, sole proprietor - all aspects of the business	Leadership
47	Attorney	Non-Technical	Everything; solo practice	Leadership
49	Nuclear medicine technologists	Technical	designer/owner Positive Train Control Operations	Non-Technical
51	All	Leadership		Non-Technical
53	distribution of food & clothing	Non-Technical	Web Administration most of the above, we are a very small business	Technical
55	QA	Technical		Leadership
57	Supply Chain	Non-Technical	Dentist	Non-Technical
59	Prefer not to say	Non-Technical	Owner, President, CEO	Non-Technical
61	Strategy/Planning I work in a juvenile detention facility.	Leadership		Leadership
63		Non-Technical		Non-Technical
65	N/A	Disqualified		Technical
67	None	Disqualified		Leadership
69	Researcher	Technical		Non-Technical

71	Not employed	Disqualified	Tax	Non-Technical
73	football coach	Non-Technical	All of the above, i am self employed and do it all.	Leadership
75	Cyber Security Manager	Technical	All, sole employee	Leadership
77	Member	Non-Technical	Founding member, volunteer of a non-profit Fair Trade organization	Non-Technical
79	Information Security National Security	Technical	Graphics Design and Product Application	Technical
81	Systems	Technical	Principal Architect	Technical
83	Product Management	Leadership	Executive Director	Leadership
85	I do it all	Leadership	Owner/Operator	Leadership
87	retired	Disqualified	everything. Single employee	Leadership
89	Account Management	Non-Technical	Owner, President and engineer	Technical
91	product safety Health and Safety	Non-Technical	Owner/Sole Proprietor	Leadership
93	Wellbeing of workers	Non-Technical	Sales, Administration, Creative, Production	Non-Technical
95	Unidentifiable	Non-Technical	Owner - all of the above	Leadership
97	None	Disqualified	President, Director of Design and Business Development	Leadership
99	Retired	Disqualified	Several: IT, Accounting, Payroll, Management - we're a nonprofit with less than 10 employees!	Technical
101	Prefer not to answer	Non-Technical	Purchasing	Non-Technical
103	Claims Lab Tech on whatever project I was assigned with	Non-Technical	I own the business and manage all projects and am sole repository of most highly sensitive information. I deal with applicants and employers.	Leadership
105		Non-Technical	Owner/Operator	Leadership
107	Automotive mechanic	Non-Technical	Sole Proprietor - I do everything.	Leadership
109	Procurement	Non-Technical	Webmaster	Technical
111	Owner/principal	Leadership	saving souls	Non-Technical
113	nurse	Non-Technical	Fundraising & Communications	Non-Technical
115	Sole proprietor	Leadership	Owner and publisher	Non-Technical
117	medical consultant	Non-Technical	sales	Non-Technical
119	Most of the above retired Health Care	Leadership	analytical Chemist	Technical
121	Provider	Non-Technical	business owner/janitor	Non-Technical
123	spiritual guidance	Non-Technical	Almost all of above because I am the owner	Leadership
125	CEO	Leadership	Church/Religious	Non-Technical
127	fabrication	Non-Technical	All of the above. I am a sole proprietor	Leadership
129	Historical Society	Non-Technical	law	Non-Technical
131	I am Retired	Disqualified	All of the above	Leadership
133	Regulatory Compliance	Technical	pastoral	Non-Technical
135	I don't know	Disqualified	Several of the above apply: Sales, Analyst, Project management, Strategy, Accounting, Customer service, Finance, IT, Research, Quality	Technical

:)

137	information security	Technical	everything	Leadership
139	Psychotherapy	Non-Technical	Pet Sitting	Non-Technical
141	Professional musician in symphony orchestra	Non-Technical	Editor	Non-Technical
143	tour guide, so basically I am the face of the company	Non-Technical	Sole proprietor and employee. As senior/solo pastor I oversee all aspects of the ministry - administrative, physical property, financial; however, we do have a bookkeeper/treasurer who takes care of the day-to-day banking & financial record keeping.	Non-Technical
145	X	Educator		Leadership
147	local church leadership Board member, Chair	Non-Technical	Owner/Leadership	Leadership
149	Health & Safety	Non-Technical	Physical Therapy Recruiter	Non-Technical
151	Husband...	Disqualified	Field services	Technical
153	retiree	Disqualified	Production of appraisal reports	Non-Technical
155	Environmental Science self-employed consultant,	Non-Technical	Vicar/Pastor	Non-Technical
157	I wear all the hats.	Leadership	Christian Faith Combination of Engineering (outdoor fiber optic networks) and project management of the construction of those designs	Non-Technical
159	Physical Security	Non-Technical		Technical
161	Installation	Non-Technical	Pet groomer	Non-Technical
163	Volunteer Coordination & Outreach	Non-Technical	Chief cook and bottle washer	Disqualified
165	Owner, Veterinarian	Leadership	Production	Non-Technical
167	Pastor	Leadership	Designer	Non-Technical
169	Chief Executive Officer	Leadership	Most of the above (small business)	Leadership
171	Administrative Medical Assistant	Non-Technical	Customer Service, Compliance and Logistics	Non-Technical
173	community services and affordable housing	Non-Technical	Other	Disqualified
175	Transportation	Non-Technical	Owner and chief DVM of a veterinary hospital	Leadership
177	Treasurer of three non-profit Community, cultural organizations	Non-Technical	To show the love of Christ	Non-Technical
179	Student	Non-Technical	manufaction i'M THE BUCK STOPS HERE PERSON SO MORE THAN ONE ABOVE APPLY TO ME.	Non-Technical
181	Judicial I am an individual real estate appraiser working from home.	Non-Technical		Leadership
183	I'm a writer and communications person for businesses of all kinds.	Non-Technical	Facilities Management	Leadership
185		Non-Technical	Everything. I am a sole practioner attorney	Non-Technical
187	Owner-do it all	Leadership	Fundraising Development	Non-Technical
189	Owner/Graphic Designer	Non-Technical	Owner, so many of the above	Leadership

191	owner, sales & customer support, only IT person (everything but manufacturing)	Non-Technical	As owner manager I wear all the hats.	Leadership
193	Corporate Executive	Leadership	Owner of tiny company, wearing many hats	Leadership
195	Social worker program production and host	Non-Technical	I'm the owner and wear many different hats.	Leadership
197	Insurance rm	Non-Technical	auto repair	Non-Technical
199	sole proprietor in legal practice - I do everything, literally	Non-Technical	editor and general manager	Leadership
201	Many above---small	Leadership	Everything..only employer and owner	Leadership
203	'mom/pop biz'	Leadership	Academic Advisor	Educator

APPENDIX G: SURVEY QUESTIONS CLASSIFICATION WITH REFERENCE NUMBER

The below table shows the numbered survey questions with the actual first phase analysis number assigned to the response as well as the purpose needed for analysis.

Response #	Question asked	Purpose
1	<i>*I agree to participate in the research study and confirm that I am at least 18 years old. I understand the purpose and nature of this study and I am participating voluntarily. I understand that I can withdraw from the study at any time, without any penalty or consequences.</i>	Consent
2. What region are you in? (Select 1)		
2	USA Europe Asia-Pacific Canada Latin America I prefer not to answer Other (please specify) Other:	Demographic (Region) Qualifying
3. What is your organization's industry or business type? (Select 1)		
3	Finance & Financial Services Advertising & Marketing Business Support & Logistics Airlines & Aerospace (including Defense) Health Care & Pharmaceuticals Education Manufacturing Government Retail & Consumer Durables Telecommunications, Technology, Internet & Electronics Nonprofit Utilities, Energy, and Extraction I prefer not to answer Other (please specify)	Demographic (Industry) Qualifying
4. How many employees are there in your organization?		
4	50 or less employees 51 to 249 employees 250 to 999 employees 1000 to 10,000 employees Over 10,000 employees I don't know I prefer not to answer Other (please specify)	Demographic (Organization size) Qualifying
5. What type of information/data does your organization collect, use, and store? (Select all that apply)		

	<i>Personally Identifiable Information (PII) – i.e., Individuals driver’s license, government id number, address, etc.</i>	
	<i>Protected Health Information (PHI) – i.e., patients medical/health records, medications, treatments, etc.</i>	
	<i>Personal data - i.e., age, gender, likes/dislikes, sexual orientation, religion, family, online social platforms info, diet, political views, pets, etc.</i>	
	<i>Employee information</i>	
	<i>Customer information or data</i>	
5	<i>Financial data – i.e., credit card data, bank accounts, etc.</i>	Poll
	<i>Student information</i>	
	<i>Data for minors, i.e., children under the age of 18 years</i>	
	<i>Your organization's intellectual property - i.e., trade secrets, procedures, designs, developed code, etc.</i>	
	<i>None</i>	
	<i>I don't know</i>	
	<i>I prefer not to answer</i>	
	<i>Other (please specify)</i>	
6. Which laws and regulations do you consider relevant to your organization? (select all that apply)		
	<i>General Data Protection Regulation (GDPR)</i>	
	<i>California Consumer Privacy Act (CCPA)</i>	
	<i>Health Insurance Portability and Accountability Act (HIPAA)</i>	
	<i>Payment Card Industry Data Security Standard (PCI DSS)</i>	
	<i>Federal Information Security Management Act (FISMA)</i>	
	<i>Family Educational Rights and Privacy Act (FERPA)</i>	
6	<i>Sarbanes–Oxley Act (SOX)</i>	Poll
	<i>BASEL II</i>	
	<i>Federal Risk and Authorization Management Program (FedRAMP)</i>	
	<i>Gramm–Leach–Bliley Act (GLBA)</i>	
	<i>Federal Financial Institutions Examination Council (FFIEC)</i>	
	<i>None</i>	
	<i>I don't know</i>	
	<i>I prefer not to answer Other (please specify)</i>	
7. How significant are the following consequences of unlawful, unauthorized, or accidental types of data incidents to your organization?		
7	<i>Reputation and brand damage – bad or embarrassing press</i>	Opinion
8	<i>Reduced revenue or customer loss</i>	Opinion
9	<i>Loss of trust on the part of interested parties</i>	Opinion
10	<i>Litigation / legal proceedings</i>	Opinion
11	<i>Deterioration of relations with employees</i>	Opinion
12	<i>Regulatory actions / sanctions or fines</i>	Opinion
13	<i>Loss of competitive advantage (for example, due to loss of intellectual property)</i>	Opinion
8. How important is improving data security in ensuring the following activities?		
14	<i>Reputation and Brand Protection</i>	Opinion
15	<i>Intellectual Property Protection</i>	Opinion
16	<i>Personal data protection</i>	Opinion
17	<i>Providing support when launching a new service</i>	Opinion
18	<i>Regulatory Compliance</i>	Opinion
19	<i>Compliance with internal policies</i>	Opinion

20	<i>Improving IT management and operations</i>	Opinion
21	<i>Increased stakeholder confidence</i>	Opinion
22	<i>Interaction with external suppliers</i>	Opinion
23	<i>Learning New and Emerging Technologies</i>	Opinion
24	<i>Assistance in mergers, acquisitions, and sales</i>	Opinion

9. Which area of your organization is your primary function or role: (Select 1)

	<i>Sales, Marketing or Business Development</i>	
	<i>Analyst</i>	
	<i>Project Management</i>	
	<i>Strategy/Planning</i>	
	<i>Accounting/Auditing</i>	
	<i>Administrative</i>	
	<i>Customer Service</i>	
	<i>Engineering</i>	
	<i>Human Resources</i>	Demographics
25	<i>Health Care Provider</i>	(Role Category)
	<i>Finance</i>	Qualifying
	<i>Educator (e.g., teacher, lecturer, professor, trainer)</i>	
	<i>Information Technology</i>	
	<i>Research and Development</i>	
	<i>Management</i>	
	<i>Consulting</i>	
	<i>Quality Assurance</i>	
	<i>Research</i>	
	<i>Student</i>	
	<i>Other (please specify)</i>	

10. What are the requirements to access your organization's systems? (Select all that apply)

	<i>Username/password</i>	
	<i>Biometric identification (fingerprint, facial recognition, other)</i>	
26	<i>Two-factor authentication</i>	Authentication
	<i>I don't know</i>	
	<i>I prefer not to answer</i>	
	<i>Other (please specify)</i>	

11. Regarding system security management, which options are true: (Select all that apply)

	<i>I am required to change my password regularly</i>	
	<i>My organization requires complex passwords (i.e., 10+ characters, must contain upper and lower alpha chars, numbers, and special characters)</i>	
	<i>My organization will not allow me to use previous passwords</i>	
27	<i>My organization has an online self-service website where I can update my password</i>	Authorization
	<i>My organization has dedicated helpdesk support if I have computer or access issues</i>	Authentication
	<i>My organization never requires me to change my password</i>	
	<i>I don't know</i>	
	<i>I prefer not to answer</i>	

12. What impact has regulatory requirements (HIPAA, PCI, GDPR, CCPA, etc.) had on the effectiveness of data security in your organization? (Select 1)

28	<p><i>Significant increase in the effectiveness of ensuring information security because of compliance with regulatory requirements</i></p> <p><i>Moderate increase in efficiency</i></p> <p><i>No change</i></p> <p><i>Information Security Efficiency decreased</i></p> <p><i>I don't know</i></p> <p><i>I prefer not to answer Other (please specify)</i></p>	Reliability
13. The following general organizational measures are in place to protect data: (Select all that apply)		
29	<p><i>The person responsible for the organization's data protection has been appointed</i></p> <p><i>Audits are done</i></p> <p><i>A policy has been developed and published to protect information (personal data)</i></p> <p><i>Developed regulations, orders, or instructions are in place</i></p> <p><i>Developed terms of reference for the creation of an information protection system are in place</i></p> <p><i>Developed technical design of information security system in place</i></p> <p><i>Information security tools have been introduced</i></p> <p><i>My organization's computer (i. e. work laptop) is updated and secured regularly</i></p> <p><i>I don't know</i></p> <p><i>I prefer not to answer</i></p> <p><i>Other (please specify)</i></p>	<p>Verification</p> <p>Authorization</p> <p>Recoverability</p>
14. If I am working on a required system at my organization and find something that isn't accurate (i.e., typo or some other incorrect information):		
30	<p><i>I can update or modify without any issue</i></p> <p><i>I need to open a support ticket or notify someone else</i></p> <p><i>There isn't anything I can do</i></p> <p><i>I don't know</i></p> <p><i>I prefer not to answer</i></p> <p><i>Other (please specify)</i></p>	Reliability
15. If a person outside my organization (member, student, customer, patient, etc.) wants or needs to update, modify, or delete personal data or information: (Select all that apply)		
31	<p><i>They can update or modify without issue</i></p> <p><i>They can delete without issue</i></p> <p><i>They must contact my organization</i></p> <p><i>It can be updated, modified, or deleted</i></p> <p><i>My organization doesn't store outside individual's information</i></p> <p><i>I don't know</i></p> <p><i>I prefer not to answer</i></p> <p><i>Other (please specify)</i></p>	<p>Verification</p> <p>Privacy</p>
16. Who in your organization oversees information security issues? (Select 1)		
32kotok1 23	<p><i>CEO, CISO, CIO, CTO, or DPO</i></p> <p><i>Other Security / Privacy officer</i></p> <p><i>Compliance Director / Director of Personal Data</i></p> <p><i>Protection Technical Director</i></p> <p><i>Head of Internal Audit</i></p> <p><i>Risk management director</i></p> <p><i>IT Specialist</i></p> <p><i>Information Security Specialist</i></p> <p><i>Network / System Administrator</i></p> <p><i>I don't know</i></p>	<p>Verification</p> <p>Privacy</p>

	<i>I prefer not to answer</i> <i>Other (please specify)</i>	
17. What if you delete an email and then you need to restore it:		
33	<i>I can easily do this</i> <i>I can contact support and they can restore it</i> <i>It's gone, and I can't restore it</i> <i>I don't know if it can be restored</i> <i>I prefer not to answer</i>	Recoverability
18. If I delete or lose a file (I.e., document, spreadsheet, etc.):		
34	<i>I can easily do this</i> <i>I can contact support and they can restore it</i> <i>It's gone, and I can't restore it</i> <i>I don't know if it can be restored</i> <i>I prefer not to answer</i>	Recoverability
19. If there is a disaster or incident affecting my "normal" tasks, I am trained by my organization, and I know what I need to do and whom I need to contact:		
35	<i>True</i> <i>False</i> <i>I prefer not to answer</i>	Recoverability Verification
20. Which of the following things are accessible to you? (Select all that apply)		
36	<i>I must be in my office to access my organization's applications</i> <i>I can access my organization email from my personal computer</i> <i>I can access organization email from my smartphone or tablet</i> <i>I have an organization laptop and remote access</i> <i>I have VPN access to my organization's systems and applications</i> <i>I can access my organization's applications with any browser over the Internet</i> <i>I don't know</i> <i>I prefer not to answer</i>	Accessibility Privacy
21. On a scale of 1 to 5, where 1 - "Strongly Disagree", 5 - "Strongly Agree", rate the following statements:		
37	<i>My organization makes data security and privacy the highest priority</i>	Opinion
38	<i>My organization, makes changing privacy and security settings easy</i>	Opinion (Ease of use) (Accessibility)

APPENDIX H: QUESTIONNAIRE SCORING RUBRIC

The below charts show the scoring matrix for each security principles based on the selected question and answer. The tallied value was then compared to all responses to determine if the principle was passed.

Authentication				
Response #	Principal	Answer Choice	Selecte d	Not Selected
10. What are the requirements to access your organization's systems? (Select all that apply)				
26	Authentication	<i>Username/password</i>	1	0
26	Authentication	<i>Biometric identification (fingerprint, facial recognition, other)</i>	8	0
26	Authentication	<i>Two-factor authentication</i>	4	0
11. Regarding system security management, which options are true: (Select all that apply)				
27	Authentication Authorization	<i>I am required to change my password regularly</i>	1	-1
27	Authentication Authorization	<i>My organization requires complex passwords (i.e., 10+ characters, must contain upper and lower alpha chars, numbers, and special characters)</i>	1	-1
27	Authentication Authorization	<i>My organization will not allow me to use previous passwords</i>	1	-1
27	Authentication Authorization	<i>My organization never requires me to change my password</i>	-1	1
27	Authentication Authorization	<i>I don't know</i>	-1	0
16. Who in your organization oversees information security issues? (Select 1)				
32	Verification	<i>I don't know</i>	-1	0

Authorization				
Response #	Principal	Answer Choice	Selecte d	Not Selected
11. Regarding system security management, which options are true: (Select all that apply)				
27	Authentication Authorization	<i>I am required to change my password regularly</i>	1	-1
27	Authentication Authorization	<i>My organization requires complex passwords (i.e., 10+ characters, must contain upper and lower alpha chars, numbers, and special characters)</i>	1	-1
27	Authentication Authorization	<i>My organization will not allow me to use previous passwords</i>	1	-1
27	Authorization	<i>My organization has an online self-service website where I can update my password</i>	1	0
27	Authorization	<i>My organization has dedicated helpdesk support if I have computer or access issues</i>	2	0
27	Authentication Authorization	<i>My organization never requires me to change my password</i>	-1	0

27	Authentication Authorization	<i>I don't know</i>	-1	0
13. The following general organizational measures are in place to protect data: (Select all that apply)				
29	Authorization	<i>Information security tools have been introduced</i>	2	0
Privacy				
Response #	Principal	Answer Choice	Selected	Not Selected
15. If a person outside my organization (member, student, customer, patient, etc.) wants or needs to update, modify, or delete personal data or information: (Select all that apply)				
31	Privacy	<i>They can update or modify without issue</i>	1	0
31	Privacy	<i>They can delete without issue</i>	1	0
31	Privacy Verification	<i>They must contact my organization</i>	1	0
31	Privacy	<i>It can be updated, modified, or deleted</i>	1	0
31	Privacy	<i>My organization doesn't store outside individual's information</i>	1	0
16. Who in your organization oversees information security issues? (Select 1)				
32	Privacy	<i>Other Security / Privacy officer</i>	1	0
32	Privacy	<i>Compliance Director / Director of Personal Data</i>	1	0
20. Which of the following things are accessible to you? (Select all that apply)				
36	Accessibility-Privacy	<i>I can access my organization's applications with any browser over the Internet</i>	-1	0

Reliability				
Response #	Principal	Answer Choice	Selected	Not Selected
12. What impact has regulatory requirements (HIPAA, PCI, GDPR, CCPA, etc.) had on the effectiveness of data security in your organization? (Select 1)				
28	Reliability	<i>Significant increase in the effectiveness of ensuring information security as a result of compliance with regulatory requirements</i>	2	0
28	Reliability	<i>Moderate increase in efficiency</i>	1	0
28	Reliability	<i>No change</i>	0	0
28	Reliability	<i>Information Security Efficiency decreased</i>	-1	0
28	Reliability	<i>I don't know</i>	0	0
28	Reliability	<i>BLANK</i>	0	0
13. The following general organizational measures are in place to protect data: (Select all that apply)				
29	Reliability	<i>My organization's computer (i.e. work laptop) is updated and secured regularly</i>	1	-1
29	Reliability	<i>I don't know</i>	-1	1
14. If I am working on a required system at my organization and find something that isn't accurate (i.e., typo or some other incorrect information):				
30	Reliability	<i>I can update or modify without any issue</i>	1	0
30	Reliability	<i>I need to open a support ticket or notify someone else</i>	1	0
30	Reliability	<i>There isn't anything I can do</i>	-2	0
30	Reliability	<i>I don't know</i>	-1	0
30	Reliability	<i>Other (please specify)</i>	1	0

30	Reliability	BLANK	0	0
----	-------------	-------	---	---

Verification				
Response #	Principal	Answer Choice	Selected	Not Selected
13. The following general organizational measures are in place to protect data: (Select all that apply)				
29	Recoverability Verification	<i>The person responsible for the organization's data protection has been appointed</i>	1	0
29	Verification	<i>Audits are done</i>	1	0
29	Verification	<i>A policy has been developed and published to protect information (personal data)</i>	1	0
29	Verification	<i>Developed regulations, orders, or instructions are in place</i>	1	0
29	Verification	<i>Developed terms of reference for the creation of an information protection system are in place</i>	1	0
29	Verification	<i>Developed technical design of information security system in place</i>	1	0
15. If a person outside my organization (member, student, customer, patient, etc.) wants or needs to update, modify, or delete personal data or information: (Select all that apply)				
31	Privacy Verification	<i>They must contact my organization</i>	1	0
31	Verification	<i>I don't know</i>	-1	0
16. Who in your organization oversees information security issues? (Select 1)				
32	Verification	<i>CEO, CISO, CIO, CTO, or DPO</i>	2	0
32	Verification	<i>Protection Technical Director</i>	2	0
32	Verification	<i>Head of Internal Audit</i>	2	0
32	Verification	<i>Risk management director</i>	2	0
32	Verification	<i>IT Specialist</i>	1	0
32	Verification	<i>Information Security Specialist Network / System</i>	1	0
32	Verification	<i>Administrator</i>	0	0
32	Verification	<i>I don't know</i>	-1	0
32	Verification	<i>BLANK</i>	0	0
19. If there is a disaster or incident affecting my "normal" tasks, I am trained by my organization, and I know what I need to do and whom I need to contact:				
35	Recoverability Verification	<i>TRUE</i>	1	0
35	Recoverability Verification	<i>FALSE</i>	-1	0
35	Recoverability Verification	<i>BLANK</i>	0	0

Recoverability				
Response #	Principal	Answer Choice	Selected	Not Selected
13. The following general organizational measures are in place to protect data: (Select all that apply)				
29	Recoverability Verification	<i>The person responsible for the organization's data protection has been appointed</i>	1	0
17. What if you delete an email and then you need to restore it:				
33	Recoverability	<i>I can easily do this</i>	1	0
33	Recoverability	<i>I can contact support and they can restore it</i>	1	0
33	Recoverability	<i>It's gone, and I can't restore it</i>	-1	0
33	Recoverability	<i>I don't know if it can be restored</i>	-1	0
33	Recoverability	<i>BLANK</i>	0	0
18. If I delete or lose a file (I.e., document, spreadsheet, etc.):				
34	Recoverability	<i>I can easily do this</i>	1	0
34	Recoverability	<i>I can contact support and they can restore it</i>	1	0
	Recoverability	<i>It's gone, and I can't restore it</i>	-2	0
34	Recoverability	<i>I don't know if it can be restored</i>	-1	0
34	Recoverability	<i>BLANK</i>	0	0
19. If there is a disaster or incident affecting my "normal" tasks, I am trained by my organization, and I know what I need to do and whom I need to contact:				
35	Recoverability Verification	<i>TRUE</i>	3	0
35	Recoverability Verification	<i>FALSE</i>	-1	0
35	Recoverability Verification	<i>BLANK</i>	0	0

Accessibility				
Response #	Principal	Answer Choice	Selected	Not Selected
20. Which of the following things are accessible to you? (Select all that apply)				
36	Accessibility	<i>I must be in my office to access my organization's applications</i>	-1	1
36	Accessibility	<i>I can access my organization email from my personal computer</i>	-1	1
36	Accessibility	<i>I can access organization email from my smartphone or tablet</i>	1	0
36	Accessibility	<i>I have an organization laptop and remote access</i>	1	0
36	Accessibility	<i>I have VPN access to my organization's systems and applications</i>	1	0
36	Accessibility Privacy	<i>I can access my organization's applications with any browser over the Internet</i>	-1	0
36	Accessibility	<i>I don't know</i>	-1	0

21. On a scale of 1 to 5, where 1 - "Strongly Disagree", 5 - "Strongly Agree", rate the following statements:

38

Accessibility *1-5 scale value*

1-5

0

APPENDIX I: GLOSSARY OF STATISTICAL TERMS

Bonferroni Correction: If one conducts a lot of correlations, some relationships will occur by chance. To mitigate this, Bonferroni correction is applied. It reduces the alpha level for the analysis, thus reducing the likelihood of making a Type I error (false positive); it is based on the number of times each variable is used.

Chi-Square Test Statistic (χ^2): Refers to the number of values used to compute a statistic. The *df* is determined from the number of groups the nominal variable has; used with χ^2 to compute the *p*-value.

Cohen's *d*: Effect size for the *t*-test; determines the strength of the differences between the matched scores. The larger the effect size, the greater the differences in the matched scores.

Correlation Coefficient (*r*): Ranges from -1 to 1; describes the strength of the relationship between the variables.

Critical Value: The minimum value at which an observed correlation coefficient is statistically significant.

Degrees of Freedom (*df*): Refers to the number of values used to compute a statistic; an *F*-test has two values for *df*: the first is determined by the number of groups being compared - 1, and the second is approximately the number of observations in the sample; used with the *F* to determine the *p*-value.

Dummy-Code: Performed in order to add a nominal or ordinal independent variable into the regression model; turns the one variable into a series of dichotomous "yes/no" variables, one for each category; one of the categories are left out of the regression as the reference group that all other categories are compared to.

Effect Size: The strength of the relationship.

***F* Ratio (*F*):** The ratio of explained variance to error variance; used with the two *df* values to determine the *p*-value.

Friedman Test: Friedman test is a non-parametric significance test for more than two dependent samples and is also known as the Friedman two-way analysis of variance; it is used as a null hypothesis test. In other words, it is used to test that there is no significant difference between the size of 'k' dependent samples and the population from which these have been drawn. The Friedman test statistic is distributed approximately as chi-square, with (k - 1) degrees of freedom.

Kurtosis: The measure of the tail behavior of a distribution. Positive kurtosis signifies a distribution is more prone to outliers, and negative kurtosis implies a distribution is less prone to outliers.

Levene's Test: Test to assess if the assumption of equality of variance is met; if significance is found, the groups differ in their spread of the dependent variable scores; this may differ from the output found from other statistical packages (such as SPSS), as Intellectus Statistics™ uses the median instead of the mean for calculations; the median tends to provide a more-robust choice that can account for non-normality.

Mann Whitney *U*: The Mann-Whitney *U* is a non-parametric test used to assess for significant differences in a scale or ordinal dependent variable by a single dichotomous independent variable. It is the non-parametric equivalent of the independent sample *t*-test. The test uses the mean ranks of the scores in each group to compute the *U* statistic, which in turn is used to compute the *p*-value (i.e., significance level). A significant result for this test suggests that the two groups have reliably different scores on the dependent variable. The Mann-Whitney *U* test assumes that the observations are independent of each other and that the dependent variable has a scale or ordinal level of measurement.

McFadden *R*²: Measures the goodness-of-fit of the model. It tends to be more conservative than *R*² values utilized in linear regression models. McFadden *R*² values of .2 or greater indicate an excellent model fit.

Mean (*M*): The average value of a scale-level variable.

Mean Rank: The average rank of the data for that group once the data is sorted and ranked.

Multicollinearity: A state of very high intercorrelations or inter-associations among a set of variables.

Non-Parametric Test: A type of statistical test that does not require the data to follow a particular distribution; typically used when assumptions of a parametric test are violated or when the data do not fit the level of measurement required by a parametric test.

Normality: Refers to the distribution of the data. The assumption is that the data follows the bell-shaped curve.

Odds Ratio (*OR*): Gives the factor increase in likelihood of the dependent variable occurring for every one unit increase in the predictor; sometimes labeled in statistical output as Exp(B).

Ordinal Data: Ordinal scales rank order the items that are being measured to indicate if they possess more, less, or the same amount of the variable being measured. An ordinal scale allows us to determine if $X > Y$, $Y > X$, or if $X = Y$.

Ordinal Data: Ordinal scales rank order the items that are being measured to indicate if they possess more, less, or the same amount of the variable being measured. An ordinal scale allows us to determine if $X > Y$, $Y > X$, or if $X = Y$.

Outlier: A data point that is abnormally distant from a set of observations.

***p*-value:** The probability of obtaining the observed results if the null hypothesis (no relationship between the independent variable(s) and dependent variable) is true; in most social science research, a result is considered statistically significant if this value is $\leq .05$.

Partial Eta Squared (η^2p): Effect size for the ANOVA and determine the groups' differences.

Percentage (%): The percentage of the frequency or count of a nominal or ordinal category.

Reference Category: Category of the dependent variable that the likelihood the other category is compared to.

Residuals: Refers to the difference between the predicted value for the dependent variable and the dependent variable's actual value.

Sample Maximum (Max): The largest numeric value in a given sample.

Sample Minimum (Min): The smallest numeric value in a given sample.

Sample Size (*n*): The frequency or count of a nominal or ordinal category.

Shapiro-Wilk Test: A test to assess if the assumption of normality is met. If statistical significance is found in this test, the data is *not* normally distributed.

Skewness: The measure of asymmetry in the distribution of a variable. Positive skewness indicates a long right tail, while negative skewness indicates a long-left tail.

Sphericity: When there are three or more repeated measurements, the variance of the differences between each pair of measurements must be equal. Sphericity is the term used to describe this measurement.

Standard Deviation (*SD*): The spread of the data around the mean of a scale variable.

Standard Error (*SE*): How much the *B* is expected to vary.

Standard Error of the Mean (*SEM*): The estimate of how far the sample mean is likely to differ from the actual population mean.

***t*-Test Statistic (*t*):** Used with the *df* to determine the *p* value.

Type I Error: Rejection of the null hypothesis when the null hypothesis is true, referred to as a false-positive result.

***U*-Test Statistic (*U*):** Used to compute the *p* value.

Unstandardized Beta (*B*): The slope of the predictor with the log-odds of the dependent variable.

Variance Inflation Factors: A measurement to assess the amount of multicollinearity present in regression analysis.

APPENDIX J: BINARY LOGISTIC REGRESSION REPORT

Binary Logistic Regression with HIPAA predicted by TotalPrinciplesPassed

Included Variables:
HIPAA and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-1.846	0.165	124.791	5.655e-29		
TotalPrinciplesPassed	0.284	0.032	77.734	1.179e-18	1.329	[1.247, 1.415]

Model Fit Statistics:
 $\chi^2 = 84.781$ on 1 df, $p = 3.333e-20$, McFadden $R^2 = 0.044$

Binary Logistic Regression with FERPA predicted by TotalPrinciplesPassed

Included Variables:
FERPA and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-3.120	0.244	164.001	1.512e-37		
TotalPrinciplesPassed	0.299	0.045	44.622	2.390e-11	1.349	[1.235, 1.472]

Model Fit Statistics:
 $\chi^2 = 50.362$ on 1 df, $p = 1.278e-12$, McFadden $R^2 = 0.039$

Binary Logistic Regression with GDPR predicted by TotalPrinciplesPassed

Included Variables:
GDPR and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-1.524	0.158	93.588	3.886e-22		
TotalPrinciplesPassed	0.229	0.031	54.387	1.646e-13	1.258	[1.183, 1.336]

Model Fit Statistics:

$\chi^2 = 57.716$ on 1 df, $p = 3.028e-14$, McFadden $R^2 = 0.029$

Binary Logistic Regression with CCPA predicted by TotalPrinciplesPassed

Included Variables:

CCPA and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-3.240	0.269	144.949	2.204e-33		
TotalPrinciplesPassed	0.253	0.050	25.968	3.470e-07	1.288	[1.168, 1.420]

Model Fit Statistics:

$\chi^2 = 28.695$ on 1 df, $p = 8.473e-08$, McFadden $R^2 = 0.027$

Binary Logistic Regression with FISMA predicted by TotalPrinciplesPassed

Included Variables:

FISMA and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-3.436	0.253	184.800	4.339e-42		
TotalPrinciplesPassed	0.379	0.046	68.696	1.149e-16	1.460	[1.335, 1.597]

Model Fit Statistics:

$\chi^2 = 81.213$ on 1 df, $p = 2.027e-19$, McFadden $R^2 = 0.060$

Binary Logistic Regression with FedRAMP predicted by TotalPrinciplesPassed

Included Variables:
FedRAMP and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-5.060	0.466	117.719	1.998e-27		
TotalPrinciplesPassed	0.439	0.080	29.872	4.614e-08	1.551	[1.325, 1.816]

Model Fit Statistics:

$\chi^2 = 37.436$ on 1 df, p = 9.444e-10, McFadden $R^2 = 0.058$

Binary Logistic Regression with PCI_DSS predicted by TotalPrinciplesPassed

Included Variables:
PCI_DSS and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-1.640	0.165	98.571	3.136e-23		
TotalPrinciplesPassed	0.182	0.032	31.790	1.717e-08	1.200	[1.126, 1.278]

Model Fit Statistics:

$\chi^2 = 33.317$ on 1 df, p = 7.828e-09, McFadden $R^2 = 0.018$

Binary Logistic Regression with SOX predicted by TotalPrinciplesPassed

Included Variables:
SOX and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-4.412	0.334	174.926	6.215e-40		

TotalPrinciplesPassed	0.474	0.058	66.582	3.356e-16	1.606	[1.433, 1.800]
-----------------------	-------	-------	--------	-----------	-------	----------------

Model Fit Statistics:

$\chi^2 = 84.518$ on 1 df, $p = 3.808e-20$, McFadden $R^2 = 0.079$

Binary Logistic Regression with BASELII predicted by TotalPrinciplesPassed

Included Variables:

BASELIII and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-5.479	0.741	54.674	1.423e-13		
TotalPrinciplesPassed	0.269	0.133	4.096	4.299e-02	1.309	[1.009, 1.699]

Model Fit Statistics:

$\chi^2 = 4.622$ on 1 df, $p = 0.032$, McFadden $R^2 = 0.020$

Binary Logistic Regression with GLBA predicted by TotalPrinciplesPassed

Included Variables:

GLBA and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-5.296	0.531	99.445	2.017e-23		
TotalPrinciplesPassed	0.426	0.091	21.707	3.176e-06	1.531	[1.280, 1.832]

Model Fit Statistics:

$\chi^2 = 27.019$ on 1 df, $p = 2.015e-07$, McFadden $R^2 = 0.052$

Binary Logistic Regression with FFIEC predicted by TotalPrinciplesPassed

Included Variables:

FFIEC and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-4.931	0.524	88.711	4.569e-21		
TotalPrinciplesPassed	0.324	0.093	12.193	4.797e-04	1.382	[1.153, 1.658]

Model Fit Statistics:

$\chi^2 = 14.194$ on 1 df, $p = 0.00016$, McFadden $R^2 = 0.032$

Binary Logistic Regression with Reg_Unknown predicted by TotalPrinciplesPassed

Included Variables:

Reg_Unknown and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-0.207	0.157	1.725	1.891e-01		
TotalPrinciplesPassed	-0.268	0.035	57.746	2.983e-14	0.765	[0.714, 0.820]

Model Fit Statistics:

$\chi^2 = 59.279$ on 1 df, $p = 1.368e-14$, McFadden $R^2 = 0.040$

Binary Logistic Regression with Reg_None predicted by TotalPrinciplesPassed

Included Variables:

Reg_None and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1468

Coefficients:

Variable	B	SE	χ^2	p	OR	95% CI
(Intercept)	-0.863	0.190	20.698	5.378e-06		
TotalPrinciplesPassed	-0.298	0.045	43.811	3.617e-11	0.743	[0.680, 0.811]

Model Fit Statistics:

$\chi^2 = 44.753$ on 1 df, $p = 2.236e-11$, McFadden $R^2 = 0.044$

APPENDIX K: POINT BISERIAL CORRELATION REPORT

Point Biserial Correlation Test for HIPAA and TotalPrinciplesPassed

Included Variables:
HIPAA and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
HIPAA-TotalPrinciplesPassed	0.235	[0.187, 0.283]	6.171e-20

Note: n = 1468;

Point Biserial Correlation Test for FISMA and TotalPrinciplesPassed

Included Variables:
FISMA and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
FISMA-TotalPrinciplesPassed	-0.224	[-0.272, -0.175]	3.415e-18

Note: n = 1468;

Point Biserial Correlation Test for SOX and TotalPrinciplesPassed

Included Variables:
SOX and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
SOX-TotalPrinciplesPassed	-0.224	[-0.272, -0.175]	3.255e-18

Note: n = 1468;

Point Biserial Correlation Test for GDPR and TotalPrinciplesPassed

Included Variables:
GDPR and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
GDPR-TotalPrinciplesPassed	-0.195	[-0.244, -0.146]	4.177e-14

Note: n = 1468;

Point Biserial Correlation Test for FERPA and TotalPrinciplesPassed

Included Variables:
FERPA and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
FERPA-TotalPrinciplesPassed	-0.179	[-0.228, -0.129]	5.594e-12

Note: n = 1468;

Point Biserial Correlation Test for FedRAMP and TotalPrinciplesPassed

Included Variables:
FedRAMP and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
FedRAMP-TotalPrinciplesPassed	-0.149	[-0.199, -0.099]	8.920e-09

Note: n = 1468;

Point Biserial Correlation Test for PCI_DSS and TotalPrinciplesPassed

Included Variables:
PCI_DSS and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
PCI_DSS-TotalPrinciplesPassed	-0.149	[-0.198, -0.098]	1.063e-08

Note: n = 1468;

Point Biserial Correlation Test for CCPA and TotalPrinciplesPassed

Included Variables:
CCPA and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
CCPA-TotalPrinciplesPassed	-0.135	[-0.185, -0.085]	1.947e-07

Note: n = 1468;

Point Biserial Correlation Test for GLBA and TotalPrinciplesPassed

Included Variables:
GLBA and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
GLBA-TotalPrinciplesPassed	-0.127	[-0.177, -0.076]	1.050e-06

Note: n = 1468;

Point Biserial Correlation Test for FFIEC and TotalPrinciplesPassed

Included Variables:
FFIEC and TotalPrinciplesPassed

Sample Size (Complete Cases):
N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
FFIEC-TotalPrinciplesPassed	-0.094	[-0.144, -0.043]	0.00033

Note: n = 1468;

Point Biserial Correlation Test for BASELII and TotalPrinciplesPassed

Included Variables:
 BASELIII and TotalPrinciplesPassed

Sample Size (Complete Cases):
 N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
BASELIII-TotalPrinciplesPassed	-0.054	[-0.105, -0.003]	0.039

Note: n = 1468;

Point Biserial Correlation Test for Reg_None and TotalPrinciplesPassed

Included Variables:
 Reg_None and TotalPrinciplesPassed

Sample Size (Complete Cases):
 N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
Reg_None-TotalPrinciplesPassed	0.178	[0.128, 0.227]	7.018e-12

Note: n = 1468;

Point Biserial Correlation Test for Reg_Unknown and TotalPrinciplesPassed

Included Variables:
 Reg_Unknown and TotalPrinciplesPassed

Sample Size (Complete Cases):
 N = 1468

Correlation Results:

Combination	r_{pb}	95% CI	p
Reg_Unknown-TotalPrinciplesPassed	0.203	[0.153, 0.252]	4.122e-15

Note: n = 1468

APPENDIX L: INDEPENDENT SAMPLES T-TEST REPORT

Independent t-Test for TotalPrinciplesPassed by HIPAA

Included Variables:

TotalPrinciplesPassed and HIPAA

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

HIPAA selected: $W = 0.896$, $p = 7.441e-19$

HIPAA not selected: $W = 0.941$, $p = 1.101e-18$

Overall: $W = 0.927$, $p = 4.569e-26$

Levene's Test:

$df_n = 1$, $df_d = 1466$, $F = 25.928$, $p = 4.004e-07$

Results:

Variable	HIPAA selected		HIPAA not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.122	1.597	4.225	1.899	9.275	6.171e-20	0.511

Note. $n = 1468$, $df = 1466.000$.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 0.707, Mean Difference = 0.896, Upper Limit = 1.086

Independent t-Test for TotalPrinciplesPassed by FERPA

Included Variables:

TotalPrinciplesPassed and FERPA

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

FERPA selected: $W = 0.869$, $p = 2.409e-13$

FERPA not selected: $W = 0.935$, $p = 1.009e-22$

Overall: $W = 0.927$, $p = 4.569e-26$

Levene's Test:

$df_n = 1$, $df_d = 1466$, $F = 21.386$, $p = 4.086e-06$

Results:

Variable	FERPA selected		FERPA not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.315	1.586	4.418	1.855	7.723	1.090e-13	0.520

Note. $n = 1468$, $df = 367.115$.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 0.669, Mean Difference = 0.897, Upper Limit = 1.126

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by FERPA

Included Variables:

TotalPrinciplesPassed and FERPA

Sample Size (Complete Cases):

$N = 1468$

Results:

$U = 186019.500$, $z = -7.002$, $p = 2.526e-12$

Medians for TotalPrinciplesPassed by FERPA

FERPA selected = 6.000 and FERPA not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by GDPR

Included Variables:

TotalPrinciplesPassed and GDPR

Sample Size (Complete Cases):

$N = 1468$

Shapiro-Wilk Test:

GDPR selected: $W = 0.899$, $p = 6.479e-19$

GDPR not selected: $W = 0.941$, $p = 2.049e-18$

Overall: $W = 0.927$, $p = 4.569e-26$

Levene's Test:

$df_n = 1$, $df_d = 1466$, $F = 8.204$, $p = 0.0042$

Results:

Variable	GDPR selected		GDPR not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.014	1.727	4.275	1.858	7.758	1.765e-14	0.412

Note. $n = 1468$, $df = 1273.141$.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 0.552, Mean Difference = 0.739, Upper Limit = 0.926

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by GDPR

Included Variables:

TotalPrinciplesPassed and GDPR

Sample Size (Complete Cases):

$N = 1468$

Results:

U = 314941.000, z = -7.580, p = 3.465e-14

Medians for TotalPrinciplesPassed by GDPR

GDPR selected = 5.000 and GDPR not selected = 4.000

Independent t-Test for TotalPrinciplesPassed by CCPA

Included Variables:

TotalPrinciplesPassed and CCPA

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

CCPA not selected: W = 0.933, p = 1.044e-23

CCPA selected: W = 0.872, p = 5.035e-11

Overall: W = 0.927, p = 4.569e-26

Levene's Test:

df_n = 1, df_d = 1466, F = 6.723, p = 0.0096

Results:

Variable	CCPA not selected		CCPA selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	4.470	1.846	5.241	1.676	-5.631	5.140e-08	0.438

Note. n = 1468, df = 233.259.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = -1.041, Mean Difference = -0.772, Upper Limit = -0.502

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by CCPA

Included Variables:

TotalPrinciplesPassed and CCPA

Sample Size (Complete Cases):

N = 1468

Results:

U = 84800.500, z = -5.363, p = 8.193e-08

Medians for TotalPrinciplesPassed by CCPA

CCPA not selected = 5.000 and CCPA selected = 6.000

Independent t-Test for TotalPrinciplesPassed by FISMA

Included Variables:

TotalPrinciplesPassed and FISMA

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

FISMA selected: $W = 0.848$, $p = 3.939e-15$

FISMA not selected: $W = 0.938$, $p = 5.168e-22$

Overall: $W = 0.927$, $p = 4.569e-26$

Levene's Test:

$df_n = 1$, $df_d = 1466$, $F = 24.226$, $p = 9.533e-07$

Results:

Variable	FISMA selected		FISMA not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.463	1.584	4.372	1.839	9.708	3.372e-20	0.636

Note. $n = 1468$, $df = 411.326$.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 0.870, Mean Difference = 1.091, Upper Limit = 1.312

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by FISMA

Included Variables:

TotalPrinciplesPassed and FISMA

Sample Size (Complete Cases):

$N = 1468$

Results:

$U = 209017.000$, $z = -8.954$, $p = 3.428e-19$

Medians for TotalPrinciplesPassed by FISMA

FISMA selected = 6.000 and FISMA not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by FedRAMP

Included Variables:

TotalPrinciplesPassed and FedRAMP

Sample Size (Complete Cases):

$N = 1468$

Shapiro-Wilk Test:

FedRAMP selected: $W = 0.790$, $p = 1.354e-09$

FedRAMP not selected: $W = 0.932$, $p = 1.454e-24$

Overall: $W = 0.927$, $p = 4.569e-26$

Levene's Test:

$df_n = 1$, $df_d = 1466$, $F = 13.848$, $p = 0.00021$

Results:

Variable	FedRAMP selected		FedRAMP not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.679	1.554	4.493	1.838	6.712	1.268e-09	0.696

Note. n = 1468, df = 97.665.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 0.835, Mean Difference = 1.185, Upper Limit = 1.535

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by FedRAMP

Included Variables:

TotalPrinciplesPassed and FedRAMP

Sample Size (Complete Cases):

N = 1468

Results:

U = 80787.500, z = -6.088, p = 1.142e-09

Medians for TotalPrinciplesPassed by FedRAMP

FedRAMP selected = 6.000 and FedRAMP not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by PCI_DSS

Included Variables:

TotalPrinciplesPassed and PCI_DSS

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

PCI DSS selected: W = 0.902, p = 1.240e-16

PCI DSS not selected: W = 0.936, p = 2.268e-20

Overall: W = 0.927, p = 4.569e-26

Levene's Test:

df_n = 1, df_d = 1466, F = 6.181, p = 0.013

Results:

Variable	PCI DSS selected		PCI DSS not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	4.967	1.726	4.377	1.866	5.924	4.384e-09	0.329

Note. n = 1468, df = 952.307.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 0.395, Mean Difference = 0.591, Upper Limit = 0.786

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by PCI_DSS

Included Variables:

TotalPrinciplesPassed and PCI_DSS

Sample Size (Complete Cases):

N = 1468

Results:

U = 273552.500, z = -5.652, p = 1.586e-08

Medians for TotalPrinciplesPassed by PCI_DSS

PCI DSS selected = 5.000 and PCI DSS not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by SOX

Included Variables:

TotalPrinciplesPassed and SOX

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

SOX selected: W = 0.833, p = 6.400e-13

SOX not selected: W = 0.936, p = 4.740e-23

Overall: W = 0.927, p = 4.569e-26

Levene's Test:

df_n = 1, df_d = 1466, F = 55.443, p = 1.634e-13

Results:

Variable	SOX selected		SOX not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.682	1.301	4.409	1.854	11.490	2.518e-25	0.795

Note. n = 1468, df = 282.314.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 1.055, Mean Difference = 1.273, Upper Limit = 1.491

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by SOX

Included Variables:

TotalPrinciplesPassed and SOX

Sample Size (Complete Cases):

N = 1468

Results:

U = 159597.000, z = -8.818, p = 1.164e-18

Medians for TotalPrinciplesPassed by SOX

SOX selected = 6.000 and SOX not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by BASELII

Included Variables:
TotalPrinciplesPassed and BASELII

Sample Size (Complete Cases):
N = 1468

Shapiro-Wilk Test:
BASEL II selected: $W = 0.849$, $p = 0.0026$
BASEL II not selected: $W = 0.928$, $p = 9.481e-26$
Overall: $W = 0.927$, $p = 4.569e-26$

Levene's Test:
 $df_n = 1$, $df_d = 1466$, $F = 0.602$, $p = 0.44$

Results:

Variable	BASEL II selected		BASEL II not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.348	1.748	4.549	1.843	2.065	0.039	0.445

Note. $n = 1468$, $df = 1466.000$.

Confidence Interval Based on $\alpha = 0.05$:
Lower Limit = 0.040, Mean Difference = 0.799, Upper Limit = 1.558

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by BASELII

Included Variables:
TotalPrinciplesPassed and BASELII

Sample Size (Complete Cases):
N = 1468

Results:
 $U = 20933.000$, $z = -2.169$, $p = 0.03$

Medians for TotalPrinciplesPassed by BASELII
BASEL II selected = 6.000 and BASEL II not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by GLBA

Included Variables:
TotalPrinciplesPassed and GLBA

Sample Size (Complete Cases):
N = 1468

Shapiro-Wilk Test:
GLBA selected: $W = 0.785$, $p = 3.289e-08$
GLBA not selected: $W = 0.931$, $p = 6.037e-25$
Overall: $W = 0.927$, $p = 4.569e-26$

Levene's Test:
 $df_n = 1$, $df_d = 1466$, $F = 7.659$, $p = 0.0057$

Results:

Variable	GLBA selected		GLBA not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.667	1.616	4.512	1.838	5.513	5.632e-07	0.667

Note. n = 1468, df = 69.387.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 0.737, Mean Difference = 1.155, Upper Limit = 1.573

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by GLBA

Included Variables:

TotalPrinciplesPassed and GLBA

Sample Size (Complete Cases):

N = 1468

Results:

U = 61177.500, z = -5.210, p = 1.889e-07

Medians for TotalPrinciplesPassed by GLBA

GLBA selected = 6.000 and GLBA not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by FFIEC

Included Variables:

TotalPrinciplesPassed and FFIEC

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

FFIEC selected: W = 0.851, p = 1.143e-05

FIEC not selected: W = 0.929, p = 2.469e-25

Overall: W = 0.927, p = 4.569e-26

Levene's Test:

df_n = 1, df_d = 1466, F = 2.774, p = 0.096

Results:

Variable	FFIEC selected		FIEC not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	5.462	1.627	4.528	1.843	3.600	0.00033	0.537

Note. n = 1468, df = 1466.000.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = 0.425, Mean Difference = 0.933, Upper Limit = 1.442

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by FFIEC

Included Variables:

TotalPrinciplesPassed and FFIEC

Sample Size (Complete Cases):

N = 1468

Results:

U = 47868.500, z = -3.731, p = 0.00019

Medians for TotalPrinciplesPassed by FFIEC

FFIEC selected = 6.000 and FIEC not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by Reg_Unknown

Included Variables:

TotalPrinciplesPassed and Reg_Unknown

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

Unknown selected: W = 0.949, p = 9.785e-09

Unknown not selected: W = 0.919, p = 1.438e-24

Overall: W = 0.927, p = 4.569e-26

Levene's Test:

df_n = 1, df_d = 1466, F = 4.188, p = 0.041

Results:

Variable	Unknown selected		Unknown not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	3.823	1.909	4.751	1.778	-7.610	1.672e-13	0.503

Note. n = 1468, df = 441.520.

Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = -1.167, Mean Difference = -0.928, Upper Limit = -0.688

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by Reg_Unknown

Included Variables:

TotalPrinciplesPassed and Reg_Unknown

Sample Size (Complete Cases):

N = 1468

Results:

U = 126309.000, z = -7.566, p = 3.837e-14

Medians for TotalPrinciplesPassed by Reg_Unknown

Unknown selected = 4.000 and Unknown not selected = 5.000

Independent t-Test for TotalPrinciplesPassed by Reg_None

Included Variables:

TotalPrinciplesPassed and Reg_None

Sample Size (Complete Cases):

N = 1468

Shapiro-Wilk Test:

None selected: $W = 0.953$, $p = 3.199e-05$ None not selected: $W = 0.922$, $p = 2.023e-25$ Overall: $W = 0.927$, $p = 4.569e-26$

Levene's Test:

 $df_n = 1$, $df_d = 1466$, $F = 1.934$, $p = 0.16$

Results:

Variable	None selected		None not selected		t	p	d
	M	SD	M	SD			
TotalPrinciplesPassed	3.625	1.902	4.676	1.804	-6.914	7.018e-12	0.567

Note. $n = 1468$, $df = 1466.000$.Confidence Interval Based on $\alpha = 0.05$:

Lower Limit = -1.349, Mean Difference = -1.051, Upper Limit = -0.753

Two-Tailed Mann Whitney U Test for TotalPrinciplesPassed by Reg_None

Included Variables:

TotalPrinciplesPassed and Reg_None

Sample Size (Complete Cases):

N = 1468

Results:

 $U = 72055.000$, $z = -6.525$, $p = 6.790e-11$

Medians for TotalPrinciplesPassed by Reg_None

None selected = 3.000 and None not selected = 5.000

APPENDIX M: SPEARMAN CORRELATION ANALYSIS REPORT 1

Spearman Correlation Test

Included Variables:

ConseqLossTrust and SecPriVal

Sample Size (Complete Cases):

N = 1324

Correlation Results:

Combination	r_s	95% CI	p
ConseqLossTrust-SecPriVal	0.258	[0.207, 0.307]	1.576e-21

Note: n = 1324;

Spearman Correlation Test

Included Variables:

ConseqReputation and SecPriVal

Sample Size (Complete Cases):

N = 1308

Correlation Results:

Combination	r_s	95% CI	p
ConseqReputation-SecPriVal	0.273	[0.222, 0.323]	7.768e-24

Note: n = 1308;

Spearman Correlation Test

Included Variables:

ConseqCustLoss and SecPriVal

Sample Size (Complete Cases):

N = 1279

Correlation Results:

Combination	r_s	95% CI	p
ConseqCustLoss-SecPriVal	0.245	[0.193, 0.296]	6.515e-19

Note: n = 1279;

Spearman Correlation Test

Included Variables:

ConseqLitigation and SecPriVal

Sample Size (Complete Cases):
N = 1300

Correlation Results:

Combination	r_s	95% CI	p
ConseqLitigation-SecPriVal	0.282	[0.231, 0.331]	3.683e-25

Note: n = 1300;

Spearman Correlation Test

Included Variables:
ConseqFines and SecPriVal

Sample Size (Complete Cases):
N = 1266

Correlation Results:

Combination	r_s	95% CI	p
ConseqFines-SecPriVal	0.337	[0.287, 0.385]	6.220e-35

Note: n = 1266;

Spearman Correlation Test

Included Variables:
ConseqEmpRelations and SecPriVal

Sample Size (Complete Cases):
N = 1222

Correlation Results:

Combination	r_s	95% CI	p
ConseqEmpRelations-SecPriVal	0.272	[0.220, 0.324]	3.091e-22

Note: n = 1222;

Spearman Correlation Test

Included Variables:
ConseqLostCompAdv and SecPriVal

Sample Size (Complete Cases):
N = 1204

Correlation Results:

Combination	r_s	95% CI	p
-------------	-------	--------	---

ConseqLostCompAdv-SecPriVal	0.278	[0.225, 0.329]	9.778e-23
-----------------------------	-------	----------------	-----------

Note: n = 1204;

APPENDIX N: SPEARMAN CORRELATION ANALYSIS REPORT 2

Spearman Correlation Test

Included Variables:

ConseqLossTrust and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1389

Correlation Results:

Combination	r_s	95% CI	p
ConseqLossTrust-TotalPrinciplesPassed	0.167	[0.116, 0.218]	3.655e-10

Note: n = 1389;

Spearman Correlation Test

Included Variables:

ConseqReputation and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1373

Correlation Results:

Combination	r_s	95% CI	p
ConseqReputation-TotalPrinciplesPassed	0.162	[0.110, 0.213]	1.509e-09

Note: n = 1373;

Spearman Correlation Test

Included Variables:

ConseqCustLoss and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1344

Correlation Results:

Combination	r_s	95% CI	p
ConseqCustLoss-TotalPrinciplesPassed	0.148	[0.096, 0.200]	4.633e-08

Note: n = 1344;

Spearman Correlation Test

Included Variables:

ConseqLitigation and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1364

Correlation Results:

Combination	r_s	95% CI	p
ConseqLitigation-TotalPrinciplesPassed	0.191	[0.139, 0.242]	1.166e-12

Note: n = 1364;

Spearman Correlation Test

Included Variables:

ConseqFines and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1324

Correlation Results:

Combination	r_s	95% CI	p
ConseqFines-TotalPrinciplesPassed	0.170	[0.117, 0.222]	5.082e-10

Note: n = 1324;

Spearman Correlation Test

Included Variables:

ConseqEmpRelations and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1283

Correlation Results:

Combination	r_s	95% CI	p
ConseqEmpRelations-TotalPrinciplesPassed	0.133	[0.079, 0.186]	1.768e-06

Note: n = 1283;

Spearman Correlation Test

Included Variables:

ConseqLostCompAdv and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1264

Correlation Results:

Combination	r_s	95% CI	p
ConseqLostCompAdv-TotalPrinciplesPassed	0.150	[0.096, 0.204]	7.866e-08

Note: n = 1264;

Spearman Correlation Test

Included Variables:

SecPriVal and TotalPrinciplesPassed

Sample Size (Complete Cases):

N = 1391

Correlation Results:

Combination	r_s	95% CI	p
SecPriVal-TotalPrinciplesPassed	0.370	[0.323, 0.414]	2.714e-46

Note: n = 1391;