ACADEMIA MILITAR

Mestrado em Guerra da Informação

Cyber Situational Awareness Dashboard for Information Security

Dissertação para a obtenção do grau de Mestre em
GUERRA DA INFORMAÇÃO

Pedro Miguel Rodrigues Tavares

Lisboa, 2019

ACADEMIA MILITAR

Mestrado em Guerra da Informação

Cyber Situational Awareness Dashboard for Information Security

Dissertação para a obtenção do grau de Mestre em
GUERRA DA INFORMAÇÃO

Orientando: Pedro Miguel Rodrigues Tavares
Orientador: Professor Doutor José Alberto de Jesus Borges

Lisboa, 2019

Dedicatória

Aos meus pais e irmã, Moura Tavares, Maria Tavares e Sofia Tavares.
À minha mulher, Andreia Lourenço, e aos meus filhos,
Rafael Tavares e Xavier Tavares.

Agradecimentos

O desenvolvimento e conclusão desta Dissertação de Mestrado não seria possível sem a colaboração direta ou indireta de vários intervenientes aos quais devo exprimir uma palavra de agradecimento.

À minha família. Aos meus pais, Moura Tavares e Maria Tavares, por sempre me terem apoiado e incentivado incondicionalmente, em especial nas minhas opções académicas. À minha irmã por toda a orientação, incentivo e apoio dedicado. À minha sobrinha, Beatriz Lourenço e cunhado, Miguel Lourenço. Aos meus sogros, João Lourenço e Luz Lourenço. À minha mulher Andreia Lourenço, pelo incentivo, paciência e ajuda adicional dedicada durante o período de realização desta dissertação. Aos meus filhos, Rafael Tavares e Xavier Tavares, que na sua tenra idade acabaram por ser privados de tempo de superior qualidade na minha companhia, fruto do esforço complementar exigido para a realização desta dissertação.

À *QlikTech International AB* pela atribuição de uma licença enquadrada no respetivo "*Academic Program*", que me permitiu a utilização do software *Qlik Sense* na implementação deste trabalho, fator essencial para a construção dos resultados da presente dissertação.

À Academia Militar e aos seus professores que me acompanharam no decorrer da realização do curso de mestrado por todo o conhecimento transferido. Aos meus colegas de mestrado.

Ao Professor Doutor José Alberto de Jesus Borges, em primeiro lugar pelo desafio lançado para a realização da presente dissertação e em segundo, por todo o apoio e orientação sem o qual não teria sido possível a realização desta dissertação.

Abstract

In an age where organizations and people live interconnected in a cyber domain and believing that with emerging new technologies the ever expanding dimension of digital data production and communications between different systems and entities greatly rises, it is increasingly necessary to build autonomous systems capable of ensuring digital security, measuring and quantifying it in context to each entity's intrinsic needs.

The aim of this work is to create a Situational Awareness Dashboard based on the identification of the state of the art regarding relevant information security metrics and the structures and architectures to implement a situational awareness program. The research design employed was a quantitative descriptive study.

The product *conceptualized, designed and implemented* in this work was developed using commercial software that is widely deployed in the business world. A survey of relevant security metrics was developed and instantiated into an original synthetic dataset, which should be scalable and shall allow from the beginning of its implementation the capacity to answer security concerns and ensure cyber situational awareness to the critical needs of an organization, therefore leveraging resilience and business continuity.

The widespread use of dashboards, such as the one that is an outcome of this work, opens up the possibility to connect predictive software for data mining and statistical analysis, while improving the security awareness through the use of structured visual information that conveys an overall operational picture consisting of the key performance indicators that are critical for specific organizational objectives or business processes.

Keywords: Cybersecurity, Situational Awareness, Analytics, Business Intelligence, Information Security Metrics, Risk Management**.**

Sumário

Numa era em que organizações e pessoas vivem interconectados num mundo cibernético e, adivinhando-se que com as novas tecnologias emergentes, a produção de dados digitais e comunicações entre diferentes sistemas e entidades aumenta consideravelmente, torna-se cada vez mais premente a disponibilização e implementação de sistemas capazes de, não só assegurar a segurança digital, como medi-la e quantifica-la face às necessidades intrínsecas de cada entidade.

O objetivo deste estudo é a criação de um *dashboard* de consciência situacional baseado na identificação do estado da arte relativamente às métricas de segurança de informação e arquiteturas que suportem a implementação de um sistema de consciência situacional. A metodologia de estudo utilizada foi descritiva com foco quantitativo.

O produto conceptualizado, projetado e implementado nesta dissertação teve como base a utilização de um software comercial, amplamente adotado no contexto empresarial. A definição de métricas foi efetuada à medida para o caso de estudo académico, sendo expansível e permitindo desde o início da sua implementação dar resposta e assegurar a consciência situacional de potenciais utilizadores face às necessidades de uma organização. A utilização do produto desenvolvido nesta dissertação permite futuras integrações com sistemas de análise preditiva que permitam melhorar a eficiência dos sistemas de segurança de informação.

Palavras-Chave: Cibersegurança, Consciência Situacional, Tratamento e Análise de Dados, Informações Empresariais, Métricas de Segurança da Informação, Gestão do Risco.

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AI	Artificial Intelligence
BI	Business Intelligence
BIA	Business Impact Analysis
C2	Command and control
C4ISR	Command, Control, Communications, Computer, Intelligence, Surveillance and Reconnaissance
CEO	Chief Executive Officer
CIA	Confidentiality, Integrity and Availability
CISO	Chief Information Security Officer
CM	Continuous Monitoring
COTS	Commercial off-the-shelf
CSA	Cyber Situational Awareness
CSADIS	Cyber Situational Awareness Dashboard for Information Security
CSDP	Common Security and Defence Policy
CVSS	Common Vulnerability Scoring system
DoD	United States Department of Defense
DoS	Denial of Service
ENISA	European Network and Information Security Agency
ETL	Extract, Transform, Load
EU	European Union
ICT	Information and Communications Technology
INFO OPS	Information operations
IS	Information Security
ISCM	Information Security Continuous Monitoring
ISMS	Information Security Management System
ISO/IEC	International Organization for Standardization and the International Electrotechnical Commission
ISRM	Information Security Risk Management
KPI	Key Performance Indicator
NIST	National Institute of Standards & Technology

OODA	Observe–Orient–Decide–Act
PDCA	Plan-Do-Check-Act
POC	Proof of Concept
PY	Previous Year
RMF	Risk Management Framework
SA	Situational Awareness
SP	Special Publication
SQL	Structured Query Language
UUID	Universally Unique Identifier
UK	United Kingdom
US	United States

Chapter 1
Introduction

1.1. Topic Research Definition

The aim of this work is to conceptualize, design and elaborate a proof of concept of a cyber security dashboard, and the associated architecture that supports it and allows the visual analysis of information gathered from heterogeneous data sources related to the theme of information security. The conceptualization of a Situational Awareness system that aggregates distinct types of data and provides a centralized source of information provides an ongoing monitoring of the risk management. This tool should provide relevant information to command and control structures (C2) and civil organization’s decision-makers. This system is intended to offer continuous monitoring of an entity's information security status, leveraging the value of the data collected from diverse sources and turning it into actionable and relevant information for the organization. It should provide in a timely manner to achieve decision-making in useful time and warranting continuously awareness, enabling a competitive superiority state in the cyber domain. Continuous monitoring is defined by the ability to maintain (near) real-time and constant awareness of the state of information security, vulnerabilities and threats to support enterprise-level risk management (Dempsey et al., 2011).

The methodology used was a quantitative descriptive study, necessarily converging in the development and implementation of a product based on an architecture that provides the capture of data from different sources, classified as an input component, an intermediate data processing layer and an information presentation layer, i.e., the output component, where relevant metrics will be implemented and presented. This last layer will be materialized in the form of a *Dashboard*, which will allow to present information to decision makers in a simple and effective manner.

1.2. Topic Research Justification

The data flows - generated by the machine-machine interaction, *Man-Man* and the binomial *Man-Machine* - are of several million records and it is expected to grow almost exponentially in the future. The data treatment, combination and analysis of these million records and, consequently, the creation of information, that is, the attribution of value to the data, proves to be a huge task and it is not feasible in a timely manner for strategic decision-making, when not carried out automatically.

By modelling the architecture for obtaining and processing data from diverse sources - such as people, processes, technology and environment - through decision support systems, combining and presenting it through data analysis applications, enables the opportunity to gain insights through internal and external forces dynamically. By combining the logic of the Observe – Orient – Decide – Act (OODA) loop, it enables timely and iterative decision-making and near earl time adjustments to potential threats and vulnerabilities, thereby leveraging competitive advantage over the adversary and efficiently managing the risk associated with information systems.

1.3. Research Objectives

The main (general) objectives of this research are to study and collect knowledge regarding information security metrics and implement a situational awareness dashboard by developing a demonstrator product which provides cyber situational awareness in line with a risk-based management that produces actionable insights on the information security state. In order to accomplish such general objectives, four specific objectives were defined:

1. To study and survey of metrics and measures for information security by revising relevant literature and established frameworks;
2. To collect metrics for the implementation of a Cyber Situational Awareness Dashboard for Information Security (CSADIS);
3. To develop processes and architecture to sustain the implementation of the CSADIS;
4. To design and implement a demonstrator product, the CSADIS proof of concept.

1.4. Research Hypothesis, Constraints and Limitations

In conjunction with the aforementioned objectives, a set of hypotheses has been defined which will guide this on and shall be verified during the present work:

1. A business intelligence (BI) or data analytics tool allow the construction and implementation of a Situational Awareness Dashboard.
2. A subset of security metrics enables a value-driven dashboard since the start of the implementation.
3. A CSADIS is a key component towards information superiority in an organization as it provides visual insights and situational awareness.

It is expected that during the elaboration of this research, several constraints can affect the development and implementation of the present work.

An implementation of a real-world BI and analytics application is a time-consuming process, usually spent between customer assessments and meetings, requirements listing, code implementation and application development. First constraint is, obviously, time. To overcome such limitation, a progressive development will be taken in consideration to provide deliverables, focused on providing valued outcomes from the start of the CSADIS implementation.

Second, is the secrecy of the topic since an information security program is, usually, a well-kept secret across organizations regarding the choices and options that are taken to implement such a critical asset. Notwithstanding several academic proposals regarding information security metrics implementation, in the real-world all comes down to time, budget, the unique needs of each organization related with its mission, strategy and goals. Therefore, a workaround solution to initiate the process is to synthetize a set of metrics that enables product deployment.

Finally, due to the secrecy of the topic, it is expected to come across an important constraint, the open availability of datasets to the public to implement a BI tool that enables the development of insights and produces a situational awareness dashboard. A work around solution in order to develop the product and reach the aim of this work, is to develop customized datasets when and if needed.

1.5. Outline

The structure of this dissertation is divided into seven chapters. This first chapter is an introduction to this work and generically introduces the research problem, the objectives of the work and desired outcome, the hypothesis as well as constraint and limitations. On the second chapter, the literature review is elaborated and several key concepts for the development of this work and the demonstrator product are presented. On the third chapter, the computational environment chosen for the product development is presented and discussed according to the requirements to sustain the situational awareness concept and a study of a situational awareness architecture is conducted. The atomic components of a situational awareness dashboard are the creation of strong and valued metrics, that is the goal of the fourth chapter, where a study of the commonly used metrics is presented and were a selection of those metrics are performed to implement the final product of this work. On the fifth chapter, the creation of the proof of concept dashboard is described. The sixth chapter contains the discussion regarding the present work and the outcome associated with the previous chapter. The seventh chapter is the conclusion of the work presenting final remarks and proposals for future work.

Chapter 2
Literature Review

2.1. Information

Information may be broadly defined as the result of data processing, in which several steps are performed, as for example extraction and treatment, oriented to a particular action, context or business goal. Data are, in turn, representative elements of facts or occurrences obtained from different sources such as agents or sensors and might be devoid from meaning if not in context.

As described by Nunes (2015), information is the "data set in context, whose form and content are appropriate for a particular use, from which it is possible to know a particular aspect or part of reality".

2.2. Cybersecurity

Several attempts have been made to define the term, but is often related with national-binding strategies or implied meanings. Generically, can be defined as the means, capabilities, processes and tools that provide safety and security in the digital domain.

It is comprised of a set of activities in the cyber domain, based on monitoring and prevention that can protect or mitigate against any cyber threat that could cause harm to people or organizations (Nunes et al., 2018). European Network and Information Security Agency (ENISA) defines cybersecurity as the security of the cyberspace, the interconnected objects that are accessible through generic digital communications networks, and the provision of capabilities to ensure security to the cyberspace domain (European Union Agency for Network and Information Security, 2015). More recently, in an attempt to provide a broader definition, ENISA (ENISA, 2017) defined Cybersecurity as the processes that "covers all aspects of prevention, forecasting; tolerance; detection; mitigation, removal, analysis and investigation of cyber incidents. Considering the different types of components of the cyber space, cybersecurity should cover the following attributes: Availability, Reliability, Safety, Confidentiality, Integrity, Maintainability, Robustness, Survivability, Resilience, Accountability, Authenticity and Non-repudiation”. Cyber security is the fundamental element in safeguarding and protecting government assets and national security, as well as protecting critical infrastructures that enhance the economy of the 21st century (CCDCOE, 2012). In cybersecurity, people, processes and technology are all entwinned, in order to achieve a secure operational environment.

Regarding Law and Policy in the European Union (EU) for cybersecurity, the Directive on Security of Network & Information Systems (NIS Directive) was the first European legislation issued on 2016 and provides member states with legal measures and tools to ensure increased cybersecurity. It defines the competent authorities to boost member state’s preparedness, the cooperation groups to support the exchange of information and operational cooperation on cybersecurity incidents. It also promotes the culture of security in organizational sectors, business and infrastructure in order to assure economic stability, compliance and readiness on cybersecurity issues (European Union, 2016).

The EU Cybersecurity Act was adopted in 2019 and defines the cybersecurity framework for Information and Communications Technology (ICT) products, services and processes, strengthens ENISA’s role on cybersecurity by amplifying its power into operational tasks and issuing a permanent mandate, besides, it complements the NIS Directive. The Cybersecurity Act standardizes the ICT certification system for services and digital solutions across the members states, increasing trust and security of the cyber domain in the EU digital market, hence, enhancing citizens awareness about the security characteristics of digital products and services and at the same time providing organizations with secure digital solutions (European Union, 2019). European Union’s CS strategies is based on five strategic priority areas (European Commission, 2013):

- Achieving cyber resilience;
- Drastically reducing cybercrime;
- Developing cyber defence capabilities and policy based on the Common Security and Defence Policy (CSDP);
- Developing the industrial and technological resources for cybersecurity;
- Establishing a coherent international cyberspace policy for the EU.

For the Portuguese National CS strategy[1], twelve objectives were defined:

- Address cybercrime;

[1] European Union National Strategies obtained at https://www.enisa.europa.eu/topics/national-cyber-security-strategies/ncss-map/national-cyber-security-strategies-interactive-map

- Balance security with privacy;
- Citizen's awareness;
- Critical Information Infrastructure Protection;
- Engage in international cooperation;
- Establish an incident response capability;
- Establish an institutionalised form of cooperation between public agencies;
- Establish baseline security requirements;
- Establish incident reporting mechanisms;
- Foster research and development;
- Organise cyber security exercises;
- Strengthen training and educational programmes.

2.3. Information Security

Information security, although on higher spectrum and strongly related with and encompassing cybersecurity, in its classic definition, is related to the confidentiality, integrity and availability of information, both at the digital and physical domains. The protection of information is of utmost importance as the disclosure, loss, modification or unavailability of an organization’s information can rise legal issues, affect profit, impact business or cause reputational damage. As such, information security is related with the security and safeguarding of information systems, such as those which process, store and transmit data across and to distinct systems with the aim of service providing. Typically, these concepts materialize through the following fundamental pillars (ISO/IEC, 2013):

Confidentiality: this pillar is related with the information being only accessible to the authorized entity. Improper access to confidential information can have devastating consequences both for state actors as also for private or organizational entities such as commerce, industry and academia. For example, improper access or disclosure of national or military security intelligence information, or improper access to innovation and development projects, as well as research patents still in progress.

Typically, the main mechanisms to ensure protection are the implementation of cryptographic and access control systems.

Common types of threats include insecure or poorly managed networks, the use of social engineering techniques to gain access to third party credentials or the use of malicious software to obtain access to those unauthorized user credentials.

Integrity: The definition of this pillar is intrinsically related to the reliability of the source of the information as well as its veracity and completeness. In addition, in this layer one can also integrate the implementation or definition of security controls and protection mechanisms at two levels, i.e. detection and prevention to ensure such integrity of information. These refer both to the origin of the information as also to its destination, as well as to the possible changes occurring to the information during its transmission through digital streams or channels.

Availability: This layer is related with the capacity or ability to guarantee access to the information, to authorized recipients only, when this capacity is not verified, it constitutes a (possible) Denial of Service (DoS) attack. Naturally, there are other factors that may make it impossible to access information and that escapes from human control such as natural disasters, but these are much more related with safety than with security, albeit possibly creating the same impact.

These three pillars of information security should ensure that the user or decision-maker can rely on the information used and its availability at their own discretion, as well as ensure that the repositories or systems in which it is stored will process the information reliably and in useful time.

These are typically the three-essential definition of the main pillars of information security. However, some authors and organizational entities understand that there is the need to extend to two other important concepts:

Authentication: It is related with the ability to secure authorization in both issuers and receivers, as well in the channels of transmission, allowing the identification and confirmation of identity. This information security control is a step towards confidentiality goal.

Non-Repudiation: The International Organization for Standardization (ISO), (2018) defines non-repudiation as "the ability to prove the occurrence of a claimed event or action and its entities of origin", as such this specifies the inability for both parties, sender and received, to deny the interaction between both.

Along with Authentication, it is one of the usual extensions to the three basic pillars of information security (Confidentiality, Integrity and Availability), (CCDCOE, 2012).

In view of this work, a situational awareness dashboard shall provide visibility, when applicable, of such information security controls related with these pillars and the state of such implementation as well as a holistic view of the information security state of an organization.

2.4. Business Intelligence, Analytics and Data Visualization

The definition of Business Intelligence (BI) is related with several transversal sets of concepts and technologies. From procedures and methodologies capable of extracting, storing, transforming and integrating data to presenting relevant and actionable information in an automatic and visual way. The end result, the information, is directed to the business users whose objective is to enhance the strategic and tactical level, promoting the informed decision taking.

A Business Intelligence system is usually composed of several architectures and techniques that enable the transformation of raw data into useful and valuable information, such architectures and techniques are usually related with the data base concept, data warehousing and data mining techniques (Wang, 2016). Through the use of a BI infrastructure, a complex analysis system can be built which provides accurate, reliable and multi-dimensional data integrated into one decision system, not only allowing an overall picture of the business state, but also allowing a self-service approach to more experienced users or management level. Thus, it provides an incremental development of insights, analysis and business discovery capabilities. As such, through the investment and use of a robust BI infrastructure and technologies, a superior business performance can be achieved resulting in the creation of a competitive advantage through high performance measurement capabilities (Peters et al., 2016).

Business Intelligence tools promote the creation of information from massive amounts of unstructured data gathered from heterogeneous sources, allowing the creation of value from disparate data across multiple systems and sources inside an organization. As Grossmann and Rinderle-Ma (2015) summarize, a BI system must provide the following features:

- Provide decision support for specific and defined goals in context with different business activities in multiple domain areas;
- The decision capabilities supported by BI has its foundation on empirical information extracted from data, based on distinct theories for information generation.
- The decision support system must have actual capabilities in information and communication technologies.

- A BI system must deliver information in a timely manner to the right people and in the appropriate form.

The earliest references date to 1958, in the (pre) digital era, are usually attributed to Hans-Peter Luhn, who was an inventor in the field of computer science at IBM. According to Luhn (1958), in his seminal work about BI and in a visionary way, he anticipated that the fast creation of information and the evolution of human and business activities would increase the growth of data-processing and decision-making needs in a timely manner in view of the enormous increase of available information. The term business referred to any collection of activities performed, whether related to commerce and industry, government, technology and science, or military. Today, as (Brooks et al., 2015) mentions, "the definition for BI has broadened to include not only technology, but also organizational and business processes".

Kimball and Ross (2013), one of the most prominent architects of Data Warehousing models and a reputable author in Business intelligence, states that the fundamental requirements of a BI system must be concerned with:

- Ease of access to information – The data and information provided by a BI system shall be simple and intuitive to the users and be readily accessible.
- Information consistency – The information, after a process of extract, transform and load (ETL) shall be fit for purpose, that is, it must be credible.
- Adaptability to change – The system must be implemented to handle changes in the future, because data and technology inevitably will change in future.
- Timely availability – Actionable information shall be accessible in an agreed timeframe. Raw data subjected to ETL processes takes time to be processed, therefore realistic expectations must be ensured.
- Information and Data security – Because organizational information is often stored in a data warehouse, the system shall be effectively managed at an information security level and all security controls shall be in place.
- Authoritative and trustworthiness – The valued outcome of a BI system is the decision support system delivered that can be made based on the analytical process of the information. As such, the right data shall be the input of such system.
- The success of a BI system is based on the acceptance of the users – Thus, the system shall be simple and fast and shall be promoted from the top business community to the end users.

Visual data analysis, which can be categorized as a by-product of BI, refers to the techniques of representing information in visual format in order to provide the user with indicative and explicit elements. According to Edward Tufte (Tufte, 2007), visual or graphical data is the most efficient way of describing, exploring, and summarizing a set of data, more effectively allowing for reasoning about quantitative information at the same time in a simple and powerful way. Thus, the application and implementation of these principles and the construction of applications for data exploration allow to create dashboards applied to the business model in each situation. These dashboards allow the user to monitor different types of data that support various operational, strategic or analytical objectives (Few, 2013). In the context of a Situational Awareness application, data visualization is of utmost importance. It provides the tools for the modulation and presentation of information in a contextualized way.

Security data is growing at an extremely fast pace, not only at a company’s internal level, but also at an external environment as a result of an interconnected world. The need for timely and objective analysis through the extraction of the maximum possible information from disparate and massive amounts of data is of paramount importance. As such, in the present work, a BI infrastructure and application will be idealized as a tool to aggregate the information in a dashboard-like application. This tool will provide insights and visual analysis which shall promote, not only, an overall overview, but also a detail analysis of each topic.

Several tools are available in the market for security analysis – such as controlling the patch deployment, log and penetration analysis - but the concept of a tool and associated architecture with distinct topics aggregated in a unique application, albeit related with information security, is an emergent subject still in discussion and investigation on multiple domains and has yet to reach to an agreement, both on the industry and defence domains.

The aim of this work is to bring a readily available, commercial, well implemented and mature Business Intelligence tool to develop a solution which would allow to extract, gather, transform and, finally, present security related information in a comprehensive and scalable way to the decision-making personnel, thus promoting the creation of value from the data.

As a target, a well-defined data model must be implemented so that scalability can be assured. As an example, the addition of predictive data shall be anticipated, thus the connection to established analytic tools – as Rattle software, for example - shall be seen as probable future addition as a middle layer for consuming and producing data for the CSA

Dashboard. Besides, a well-structured data model shall centralize the information as to avoid the creation of data islands or information silos, focusing on the availability of information to the relevant users and business units.

Using the principles of an Information Security Management System (ISMS) implementation and the intelligence cycle, and based on past work experience, it is understood that a SA BI solution can be implemented to handle the provided data and bring forth relevant information, by extracting value and providing return of investment, promoting accountability, fulfilment of governance and leveraging information security situational awareness.

2.5. Metrics, Measures and Security Metrics

Metrics and measures are used every day in business processes interchangeably and often attributed to incorrect contexts or definitions. As such, there is the need to identify and explain the key differences between measures, metrics and KPI's and how state of the art is defining them. Metrics provide progression overview towards an initially defined objective related with a business activity. National Institute of Standards and Technology (NIST) defines a measure as a mean to measure more concrete or objective attributes, whereas metrics are used for more abstract, higher-level or subjective attributes (Black et al., 2008). The Data Warehousing Institute[2] (TDWI) (Russom et al., 2010) defines twelve intrinsic basic principles of effective measures:

- Strategic – Metrics shall focus on objectives and goals, in that sense, they shall work backwards, as outcomes are defined, a metric should be created to measure the achievement of the goal.
- Simple – Performance metrics shall be well defined; the calculation methods and targets shall be understood by the users.
- Accountability – All performance metrics shall have an owner who is responsible for its outcomes.
- Actionable – Metrics shall provide corrective actions, if a metric measure something that cannot be changed to meet a goal, it is useless.

[2] Former *The Data Warehousing Institute* now called *Transforming Data With Intelligence* (Russom et al., 2010).

- Timely – A metric shall be provided in a timely fashion as to allow the opportunity to apply corrective measures and to change course of actions. As such, frequent updates of data shall be executed.
- Referenceable – A metric shall be related with its metadata in order to create a trustworthy feeling in the users. A user shall know the origin of data, last update date/time and other relevant properties.
- Accuracy – Metrics shall be based on reliable data and on strong data transformation processes and cleansing methods.
- Correlation – Organizations shall evaluate performance metrics to ensure they drive to the desired goals.
- Game-proof – Performance metrics shall be tested and audited in order to prevent the users to circumvent them.
- Aligned – Metrics shall be aligned with organizational goals and objectives and prevent from unintentionally created sub-optimization.
- Standardized – Users shall agree on the defined metrics and must be reproducible under similar circumstances, even on distinct tools, to prevent from inconsistent performance dashboards.
- Relevant – A metric has a life cycle, over time the impact will start to diminish, therefore it must be object of evaluation to the effectiveness provided.

Information security metrics follow the same guidelines and principles, NIST definition for Security metrics is widely used and it states that, (Bowen et al., 2006) "Metrics are tools designed to improve performance and accountability through the collection, analysis, and reporting of relevant performance related data. Information security metrics monitor the accomplishment of goals and objectives by quantifying the implementation level of security controls and the efficiency and effectiveness of the controls, by analysing the adequacy of security activities, and by identifying possible improvement actions.", also, Barabanov, et al. (2019) define metrics of information security as, "measurement standards that facilitate decision making by quantifying relevant data, where measurement refers to the process by which they are obtained. A distinction between a metric and a measurement can also be drawn, where the latter quantifies only a single dimension of the object of measurement that does not hold value (facilitate decision making) in itself, while the former is derived from two or more of the latter to demonstrate an important correlation that can aid a decision”.

Voeller, et. al. (2008), define metrics as “tools to facilitate decision making and improve performance and accountability. Measures are quantifiable, observable, and objective data supporting metrics (...) Effective security metrics should be used to identify weaknesses, determine trends to better utilize security resources, and judge the success or failure of implemented security solutions, furthermore Voeller, et al. (2008), distinguish between metrics and measures as a measure being “a concrete, objective attribute, such as the percentage of systems within an organization that are fully patched, the length of time between the release of a patch and its installation on a system, or the level of access to a system that a vulnerability in the system could provide. A metric is an abstract, somewhat subjective attribute, such as how well an organization’s systems are secured against external threats or how effective the organization’s incident response team is”.

As Alberts *et al.*, (2001) defines, a “metric is a standard of measurement: measuring specifically the dimensions, capacity, quantity, or other characteristic of an attribute so that comparisons can be made”. The development of metrics enables the establishment of a relationship between information systems and associated security activities, thus demonstrating the value of information security for an organization (Elizabeth Chew et al., 2008). As such, the goal is to define useful metrics that are effectively used in security and risk management and that provide efficient decision making, not constituting a waste of resources, for this it is necessary to define goals in a hierarchical way and underlying metrics which answer the key questions for each of the objectives (Freund & Jones, 2015).

The implementation of an Information Security (IS) metric program might become a tedious and long process and a sponsor program from (top) management shall be seek when designing and adopting such an endeavour. To make the process fast, the implementation of such a program can be stranded with an organizational framework already in place as to provide an agile or incremental implementation. In seeking and designing such metrics, there is the need to adopt a strategy for designing metrics that bring value and not just a playbook of metrics. As such, Payne (2006) defines seven-step methodology to guide in the development of a metric program that bring forth usefulness and improvements in the overall security program:

- Define the metrics program goal(s) and objectives;
- Decide which metrics to generate;
- Develop strategies for generating the metrics;

- Establish benchmarks and targets;
- Determine how the metrics will be reported;
- Create an action plan and act on it, and;
- Establish a formal program review/refinement cycle.

In summary, the fundamental objectives of implementing IS metrics are to evaluate the efficiency of implementing controls, extend the implementation of security requirements, facilitate the implementation and evaluation of performance in terms of the overall risk management of the organization, and provide decision making by decision-makers (ISO/IEC, 2018). In short, they provide the ability to evaluate performance, optimize the level of protection of entities, establishing benchmarks that provide levels of monitoring, evaluation and optimization that allow the integration of business processes together (Tashi & Ghernaouti-Hélie, 2007).

The main goal of a strong security metrics system is to provide business continuity, preventing or minimizing incidents and thus, reducing the potential underlaying impact in the organization (Kott, et al. 2014).

Initially, it is necessary to identify the important metrics to be implemented for an organization, the essential issues are to interpret business model’s needs and objectives, to consider references and competitors, to use audit results to identify risk and consequently to decide the type of metric, that is, whether it is operational, management or governance related (Michael Hoehl, 2010). In sum, to identify the value of using objective metrics to take advantage of an information security plan (Rathbun, 2009a). Through these metrics, it will be possible to remedy possible vulnerabilities and even predict threats, providing information protection and, ultimately, asset protection.

According to the above definitions, it is safe to assume that both measures and metrics are performance indicators which provide qualitative and quantitative outcomes. However a measure is usually the result in the form of a numerical observation of a single event in a single point of time, whereas a metric is a result based in one or multiple measures where some calculation is applied in some given business or organizational context to provide an observation that allows monitoring and tracking of success or failure of a business goal.

A Key Performance Indicator (KPI) is a performance metric that reflect strategic drivers for the organization, therefore providing guidance towards a defined objective and allowing the understanding of how an organization is performing towards a strategic goal.

Thus, KPI's create focus on improvements and continual monitoring of key goals performance. As such, KPI shall be carefully defined on key activities that provide the most valued outcomes to the organization, and so, providing guidance towards a strategic goal with focus on value.

Security metrics shall be tailor-made for a specific organization in order to provide strong sustainability to goals and key performance indicators. Consequently, a security metric system shall be reviewed and aligned with specific organizational needs (Kott et al., 2014).

Information Security metrics follow the same rules as general metrics in business domains, and as such, the same goals, to supply the business with efficiency, give direction, appoint accountability and provide return on investment, therefore justifying the implementation of an IS program.

2.6. Situational Awareness

Multiple definitions have been raised in an attempt to define Situational Awareness, more specifically in the Cyber domain.

It is a subject increasingly studied in the recent years and as the dependency on the technology and in the digital domain for every day operations grows, so does grow the need to analyse, comprehend and to predict situations on the cyber domain.

This fast-passed growth of technology puts pressure on the needs for accurate information and for the possibility to correlate and anticipate changes on the cyber landscape of an organization or a state actor. Hence, the decision makers must be aware of the current situation as an enabler for valued informed decisions, be it at the defence or security level.

One of the most cited and first ever definition for SA, is the one by Endsley (1995), and states that SA is the "perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future".

Situational Awareness can be seen as three-phase process in which perception must provide awareness about the current status and dynamics of the elements, comprehension of the situation through information analysis and finally, projection through prediction obtained from information gathered from the first two elements. Thereby, providing accurate information and knowledge in a data triage system way to enable answers for four central questions in CSA (Liu, et al. 2017):

- What has happened?

- What is the impact?
- Why did it happen?
- What should I do?

In an attempt to secure information, most organizations have deployed multiple cyber security tools to protect their assets such as firewalls, anti-virus, intrusion detection, that create alarms but an overall and holistic overview of the information security domain is most of the time lacking (Han et al., 2019), as such, a multi data analysis tool capable of creating awareness on global level is a desired tool to be implemented.

Tianfield (2016) stated that, "awareness is contextual understanding built on intelligence, and situation(al) awareness is to get a grasp of what is happening and how it had evolved in the recent time and how it might trend away in the near future", through appropriate evaluation and inference mechanisms that provide understanding of the situation and related dynamics of the circumstances, therefore providing analysis and situational visualization of the surrounding environment for the purpose of effective decision-making in a timely manner (Tianfield, 2016).

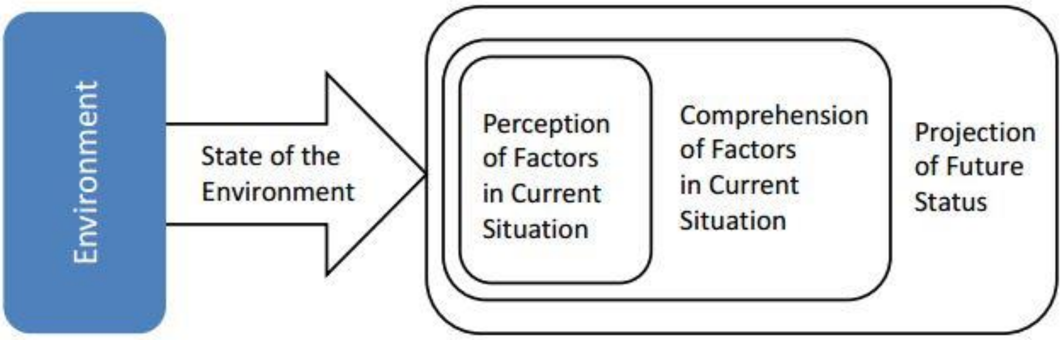

Environment
State of the Environment
Perception of Factors in Current Situation
Comprehension of Factors in Current Situation
Projection of Future Status

Figure 1 - Three layer of situational awareness – source: (Tianfield, 2016)

The goal of implementing an architecture and a model of Situational Awareness is to provide an organization with a continuous monitoring capability through the aggregation of diverse and dispersed information, from distinct systems and different implementations, allowing its analysis in a centralized manner in such a way that provides the attribution of scoring models and risk management goals (Mell, et al. 2012).

It is now accepted that the mere identification of events of potential cyber-attacks is inefficient, it is necessary to go beyond this vision and implement other types of solutions, which include the correlation of events, the visualization of data and creation of information

(European Commission, 2019) based on not only possible attacks and vulnerabilities, but also other competences, in particular the human scope, such as adherence to an information security awareness or efficient training of employees on these issues. Therefore, by combining information from different sources and natures it is possible to obtain insights into the current situation (Franke & Brynielsson, 2014).

For this, it is necessary to develop an unidirectional architecture, which transits the relevant information coming from the operational area to a section of data aggregation and availability of analysis tools, the latter should combine data not only from the operational scope, as well as monitoring physical security and business systems, granting a comprehensive situational awareness analysis (Jim McCarthy et al., 2017).

In summary, it is necessary to have a holistic view about the organization and its business components and that these are effectively defined as collaborative organizations and endowed with efficient technological capacity, ensuring the success of the implementation of knowledge monitoring or situational awareness (Matthews, et al. 2016).

Based on the above definitions, most authors agree that SA is, generically, based on three-layer model, being the first the data acquiring level from which an heterogenous multitude of data sources is to be extracted - for example sensors, agents, environmental sources, human factors - for the desired metrics and measurements that provide awareness at the subsequent levels. The second layer, or the construction of information where the data is subjected to the process of data transformation, that is, cleansed, fused, and given the disparity of data sources, it will most probably be modulated into a multi granularity dimensional model which feeds the data visualization tool. At the third layer, the analysis and comprehension layer, it is where assumptions and projections can be constructed, correlated or being a subject of AI tools, as a consequence, this layer is where the presentation of actionable information happens. Actionable information is constructed based on five key properties (ENISA, 2014):

- **Relevance**: information is considered relevant when it is addressed to the responsible entity, allowing the delivery of correct information to where actions can be taken.
- **Timeliness**: Distinct type and amounts of data might take non-identical time to be transformed into actionable information, however all users shall be aware of the idiosyncrasy of this process and shall not introduce unnecessary delays.

- **Accuracy**: Data shall be processed into accurate information by obeying to a set of pre-defined steps (data transparency sources, data cleansing, etc) so that the recipient is able to consume it immediately.
- **Completeness**: whenever possible all data shall be presented in the most complete form, both for the producer and to the consumer, allowing the understanding of the information provided.
- **Ingestibility**: information shall be provided in a straightforward way, to allow the direct use of actionable information as fast as possible.

Cyber situational awareness is closely related with the definition of an information security continuous monitoring like the one described in NIST 800-137s. As defined by Dempsey *et al.*(2011), an "information security continuous monitoring is defined as maintaining ongoing awareness of information security, vulnerabilities, and threats to support organizational risk management decisions", established by providing the means, *id Est*, using the metrics, to evaluate the effectiveness of the implemented security controls through the readily available data provided by these controls. In order to implement such an information security framework, and consequently, a metric's program, it is required to obtain an interdepartmental or transversal involvement. It must start at the top management level, by obtaining senior management's sponsorship, governance definition and strategic vision, to operational individuals who develop and implement business processes, in short the continuous monitoring program should be implemented with the following strategic factors in mind (Dempsey et al., 2011):

- A concise and factual understanding of the organizational risk management (tolerance and appetite for risk), therefore creating a consistent process prioritization and risk management across the entire organization;
- Creation of a metrics program that provides insightful indicators of the security status across the organization;
- Providing continuous effectiveness of the implemented security controls;
- Verification of compliance of the information security requirements against regulations and standards;
- Visibility of the status of all the organization's IT assets.
- Provide change control of the organization's systems and operational environments;
- Maintaining awareness of threats and vulnerabilities;

By aligning an ISCM framework to such factors, the organization's information security capabilities will enhance and mature over time to better respond to the threat and vulnerability landscape and act accordingly to the organization's risk management, in other words, to act from the organization's perspective in whether to accept, transfer, reject or mitigate risk (Dempsey et al., 2011).

Thus, SA is based upon a collection of relevant data captured and provided strategic information and shall be constructed having in mind defined goals and strategies. To achieve this, there is the need to take into account the organization's reality, dimension and information security needs, supported by a strong metrics program. Such program shall start simple and efficient, but providing not only a holistic overview but factual and objective information in accordance with the organization's information security strategy to support organizational risk management and informed decision-taking. With today's technology and tools that supports automated data capture and with the use of data analysis tools, near real-time security monitoring applications can be deployed to ensure risk-based decision making (Dempsey et al., 2011).

2.6.1. OODA loop

The OODA loop (Observe–Orient–Decide–Act) was conceptualized by John Richard Boyd (Boyd, 1987), US Air Force pilot and military strategist. Applying it to the cyber-security and information security context, this cycle is essentially characterized by the ability to enable fast decision-making by observing and interpreting multiple external and internal variables to anticipate the competitor in a continuous monitoring system.

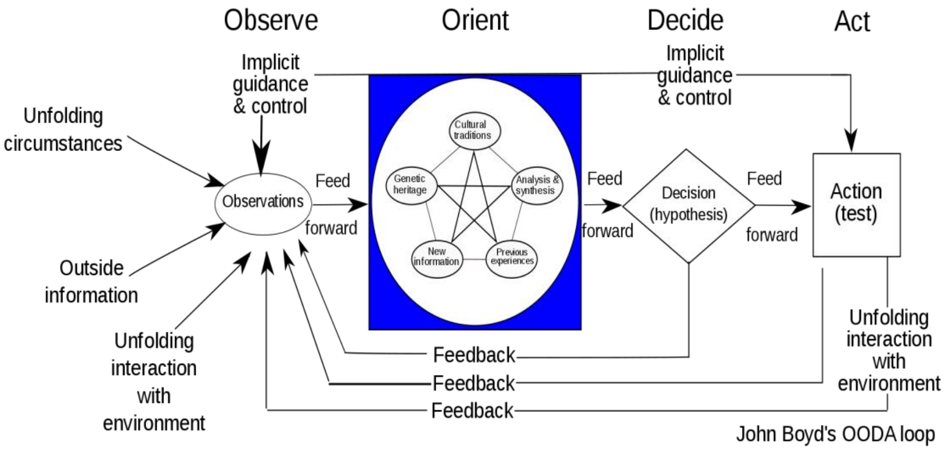


Figure 2 - OODA loop (adapted from Wikipedia, 2019)

As explained by Nunes (2015), this cycle, applied to the context of information security, allows for "shortening the observation and orientation processes by allowing sensors and decisions to be integrated into the same network and producing ever faster decision cycles. The commander/manager will now essentially have to decide and act, based on available information".

As such, in the context of this work, the relationship of a situational awareness dashboard and the OODA loop is that the applicability of such principle allows an ever-evolving development and/or update of the relevant information necessities in context with the observations of, not only the internal, but also with the external environment. Thus, such dashboard can start in a simple way, but at the same time ensuring the creation of value, and evolve as needed. This means that, besides adding metrics and data as needed, if some metric fails to produce the desired outcome, then they should be adapted or eliminated.

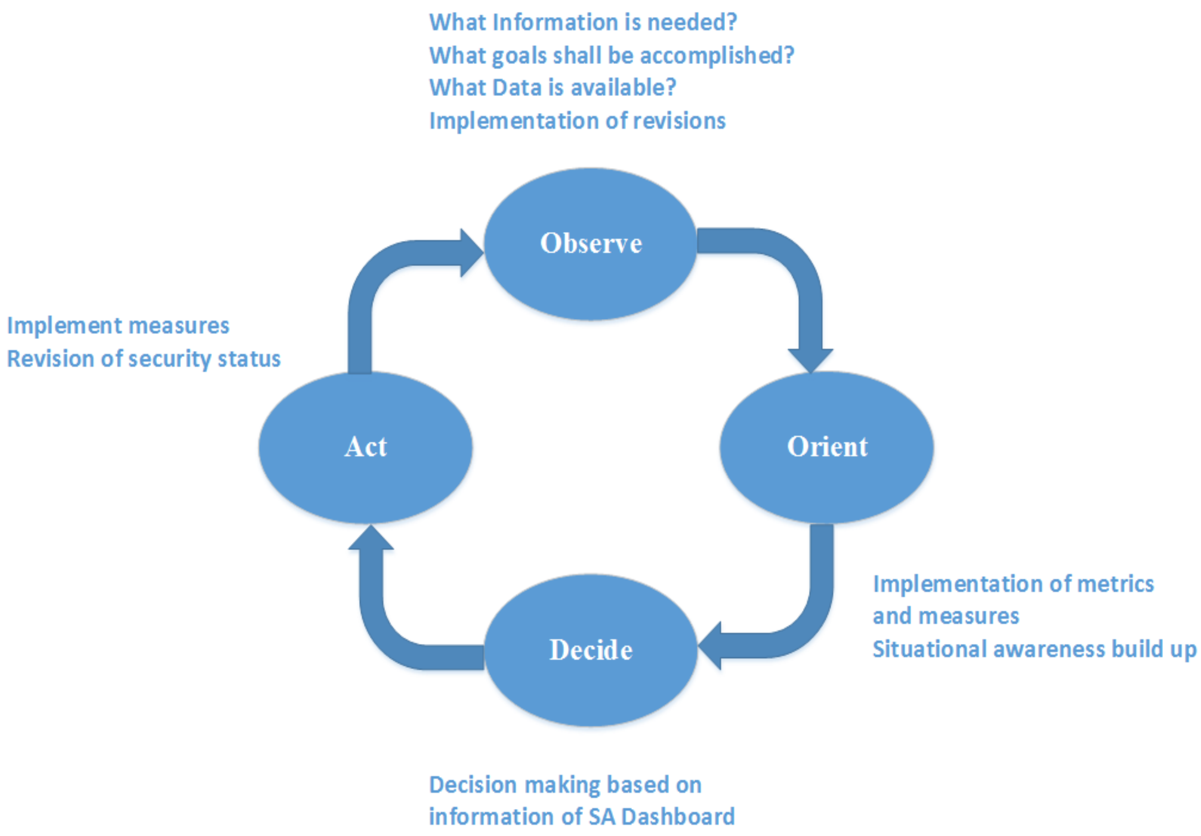

What Information is needed?
What goals shall be accomplished?
What Data is available?
Implementation of revisions
Observe
Implement measures
Revision of security status
Act
Orient
Implementation of metrics
and measures
Situational awareness build up
Decide
Decision making based on
information of SA Dashboard

Figure 3 - Proposed OODA Loop in SA Dashboard Context

As such, an organization can identify improvement opportunities, prioritize such opportunities, plan actions to materialize them and be able to monitor and evaluate such changes. With this continuous observational and adaptation process, continuous improvements can be achieved, and so the organization will be able to react and adapt in a fast-changing environment to the threat landscape (multiple attack vectors and surfaces) and to newer requirements or needs that can rise and move forward into an information superiority state.

2.6.2. Information Superiority

Alberts, Garstka and Stein (1999), defines Information superiority as “the operational advantage derived from the ability to collect, process, and disseminate an uninterrupted flow of information while exploiting or denying an adversary’s ability to do the same”.

Therefore, "The superiority of information is a state that is achieved when a competitive advantage is derived from the ability to exploit information from a superior position", i.e. when there is a perception, capacity to assimilate and treat data and information in an integral perspective and of maximum capacity” (Alberts et al., 2000).

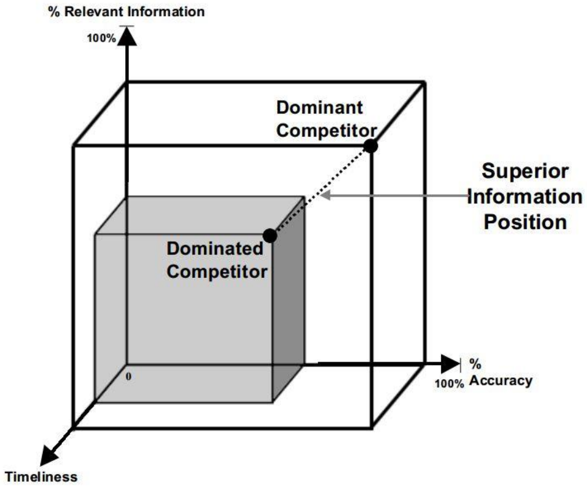


Figure 4 - Information superiority – source: (Alberts et al., 2000)

United Kingdom’s (UK) military doctrine (Note et al., 2013), albeit claiming that “there is no consensus across Defence as to what information superiority is, in its most fundamental form, nor is there a single well-understood and endorsed doctrinal definition (...)” provides an approximate definition of what information superiority is by describing it as “the competitive advantage gained through the continuous, directed and adaptive employment of relevant information principles, capabilities and behaviours”.

Information superiority is a state that supports decision-making mechanisms as an enduring principle, it supports the decision-making process providing it with information and intelligence, thus it gives information advantage to enable decision-maker to make effective decisions in an ever-changing environment in a timely manner. And in consequence, providing the ability to develop insight, foresight and understanding of the circumstances, that is, heightened awareness, enabling decision superiority through the understanding of the environment and using pattern-recognition that provides capabilities to detect changes. Information superiority is a state that is related with the overview of the real picture enabled by the capacity of situational awareness. (Note et al., 2013).

The correct decision-making process has its own underlayer on the sum of relevant information. Ambiguity and errors are important factors that might rise in an overwhelming abundance of data sources and usually has its roots on 3 factors (Laudy et al., 2006):

- Perception – lack of understanding of the context of the information or the strategy for the outcomes to be achieved.
- Comprehension – an analyst might misunderstand the gathered information of might not be prone or fully aware to some mental models.
- Projection – Biased assumptions might contribute to a wrong prediction of an incoming status.

Information Superiority is often used as a goal in a Command, Control, Communications, Computer, Intelligence, Surveillance and Reconnaissance (C4ISR) system, that is, in a network centric warfare context. In the present work the focus shall be the intrinsic properties a Situational Awareness (SA) has as an enabler to support decision in a Command and Control (C2) and Intelligence cycle, and thus being a key component of the information superiority concept and at the same time an important component in info ops[3].

In the context of the objective of this work, the SA dashboard and its implementation architecture, it is of paramount importance to stress the need of well-defined goals, with focus on valued outcomes, starting as a top down approach, by adapting to an already established or used framework in the organization and to build agreed models that supply guidance, so that all users work in one common direction.

[3] Info Ops is defined by (US Army, 1996), FM 100-6, as "continuous military operations within the military information environment that enable, enhance, and protect the friendly force's ability to collect, process, and act on information to achieve an advantage across the full range of military operations; information operations include interacting with the global information environment and exploiting or denying an adversary's information and decision capabilities".

With the implementation of a system like a situational awareness dashboard, an organization is a step forward into reducing the gap between the available data and the needed information, therefore moving towards an information superiority state.

2.7. Risk Management

As defined in (ISO, 2018), risk is the effect of uncertainty on objectives, an unexpected deviation from goals at different levels, be it strategic, tactic, organizational-wide or operational. The same document, defines principles and guidelines that implement the framework for risk management irrespective of an organization's size, activity and/or industry, and is also applicable to a broad variety of risk categories, for example, technological, financial, legal, healthcare. The ISO 3100 that treats Risk Management, provides a high-level guide that shall be tailored to an organization needs and it is based on three main principles that allow for effectively manage risk:

- Principles;
- Framework;
- Processes;

Risk perception may differ amongst distinct interlocutors and from real data to the desired goals (Borges, 2015), as such, a strong implementation of a situational awareness based on the risk assessment provides the means to normalize, understand and manage the risk associated with the objectives of an organization by means of its information security program and business continuity. Risk management and risk assessment triggers the need to build a centralized system to evaluate the current information security situation and to provide a continuous monitoring environment. In essence, this can be achieved by building a robust situational awareness dashboard in which a centralized gathering of disparate and heterogenous data is collected and treated, and where a collection of metrics and agreed business rules presents the relevant information based on the risk assessment previously elaborated, providing an organization-wide normalization of the risk in which all interlocutors access and monitors the information in the same context. The SA dashboard goal is to enable the means by which decision-making can be made based on risk analysis and treatment in an agreed and defined organizational context. With the increasing importance and associated specificities of the information systems, several risk management frameworks and guidelines specific to IS have been developed to address the necessities of such systems.

As the goal of a situational awareness is to provide information on the risk management of an organization's information systems, in the subsections below it is presented a general overview of the most renowned risk-management frameworks for information security, which addresses the particularity of attack surfaces and attack vectors associated with the information security and cyber domain and provides guidelines for the continuous monitoring of such events.

2.7.1. National Institute of Standards and Technology Special Publications

NIST SP 800-39 – Managing Information Security Risk

National Institute of Standards & Technology's[4] SP 800-39, which treats Managing Information Security Risk: Organization, Mission, and Information System View, provides a structured and adaptable proposal to an organization on how to deal and manage with information security risk related with operation and usage of information systems. It addresses the need to provide to a broad audience the authority and responsibility of defining and manage information security risk by emphasising the need of having top management owner and sponsorship, by establishing governance models for risk management, encouraging the creation of a link between business processes and associated risk and provides an holistic view of how can a risk associated with an information system process can jeopardize an entire organization and, thus, how this can affect the whole business success (NIST 800-39, 2011). It is based on a multi-level risk management approach which is strongly attached to the organizational architecture, more precisely as a three-tier model, as shown on figure 5.

[4] NIST stands for the National Institute of Standards and Technology, a non-regulatory agency of the United Stated Department of Commerce and it is responsible for the elaboration of measurements, standards and principles which shall guide and promote both innovation and competitiveness by which government – and other private industrial or commerce sectors – shall govern.

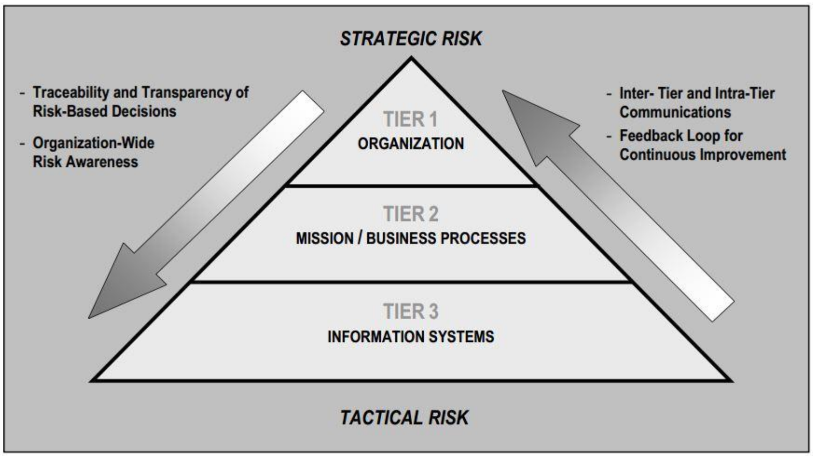

STRATEGIC RISK
- Traceability and Transparency of Risk-Based Decisions
- Organization-Wide Risk Awareness
TIER 1
ORGANIZATION
- Inter- Tier and Intra-Tier Communications
- Feedback Loop for Continuous Improvement
TIER 2
MISSION / BUSINESS PROCESSES
TIER 3
INFORMATION SYSTEMS
TACTICAL RISK

Figure 5 - Three tier model for organization-wide risk management, source: (NIST 800-39, 2011)

Tier one which is related with the organization, contextualizes the risk activities to the business goals and strategy, assigns responsibilities and inter-operability among all responsible stakeholders. At tier two, which is at mission and business process level, is related to the need to define the relevant business processes and resources that promote the organizational goals and assign a risk assessment, by defining threats, vulnerabilities and probable impact. Lastly, the tier three that deals with the system level or operational environment, deals with the risk context, risk decisions and risk activities defined at tier one and two. This is done by categorizing the organization's information systems that sustain the business processes and defining security controls to the information systems by implementing, managing and controlling of such security measures (NIST 800-39, 2011). Thus, the aim of the document is to provide the means to deal from a strategic to a tactical implementation, covering an organization-wide risk management.

NIST SP 800-37 – Risk Management Framework for Information Systems and Organizations

National Institute of Standards & Technology's SP 800-37 which treats Risk Management Framework for Information Systems and Organizations, defines a holistic risk management process that provides processes for each step of the risk management framework directed to the development lifecycle. It consists of a six-step risk management framework whose purpose is to guide the implementation of efficient and cost-effective cyber security implementation process aligned with business goals sustained by information systems, that promotes the development of security and privacy capabilities into those system, during the development and maintenance lifecycle, and aligned with the risk evaluation (NIST 800-37, 2018). An overall overview of the seven-step risk management is presented in figure 6.

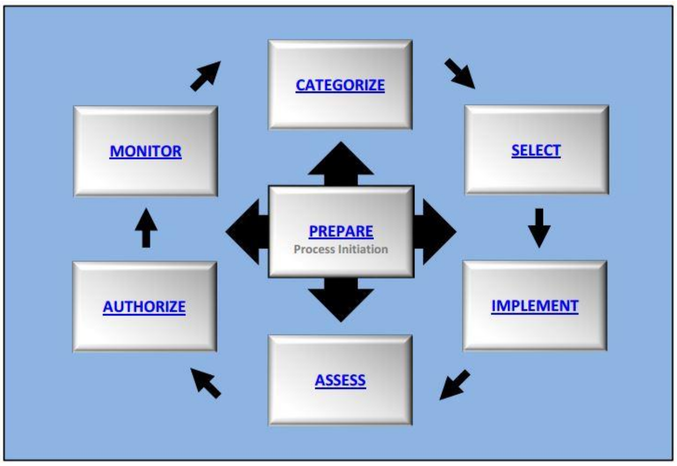


Figure 6- Six step risk management implementation, source: (NIST 800-37, 2018)

The first step, Prepare, promotes an establishment of context and priorities from security and privacy from an organization-wide to operational level. The Categorize step, is related with the assessment of the threat, vulnerability and loss impact on the information asset, that is, it concerns to the evaluation and determination of the criticality of the information value according to the impact on the organization business and function goals. The third step, Select, defines the need to select and tailor, based on the risk assessment and risk tolerance, an initial set of security controls against the organization needs, that provides a reduction of risk into an organizational-wide acceptable level that protects the confidentiality, integrity and availability that meets organizational defined requirements.

The Implement step, draws the plan on how the previous selected controls are implemented in the operational and development phase by setting the means of how to plan and defines policies and configuring settings of the operational controls. At step five, Assess, is related to the effectiveness of controls in place, its efficiency and implementation correctness, that is, if they operate and deliver the outcomes as intended in accordance to the security requirements and information assurance. The sixth step, Authorize, evaluates the security controls output in terms of risk acceptance of the implemented security controls and its threats and vulnerabilities in an integrated organizational and assets view of the desired outcomes.

Monitor, the last step, strongly related with the ongoing and continual monitoring of controls and change management of the information systems, providing information about their effectiveness in obtaining information about signs of attacks and changes that influences information security controls and their effectiveness (NIST, 2018).

´

NIST SP 800-30 – Guide for Conducting Risk Assessments

The National Institute of Standards & Technology’s SP 800-30 Guide for conducting risk assessments provides the tools to assess risk in an overall risk assessment to all three tiers of the cyber domain in accordance to Special Publication 800-39 and the steps in risk management framework described above – SP 800-37 - on a continuous monitoring basis by describing the risk management and assessment processes and, also, the means to communicate the risk results to relevant key-maker management personnel (NIST, 2012). It identifies the risk model, by defining the concept of threats, threat shifting, vulnerabilities and predisposing conditions, likelihood and impact from a threat as illustrated on figure 7.

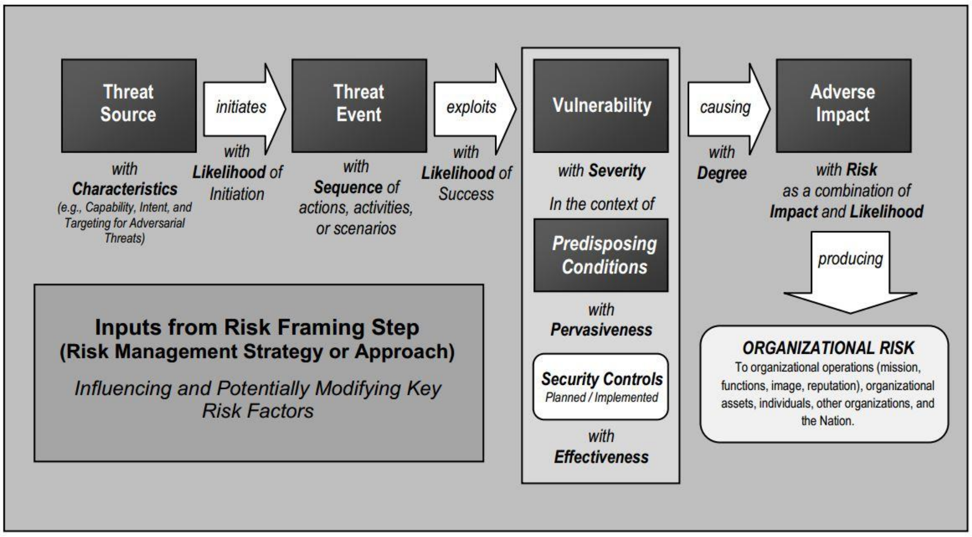
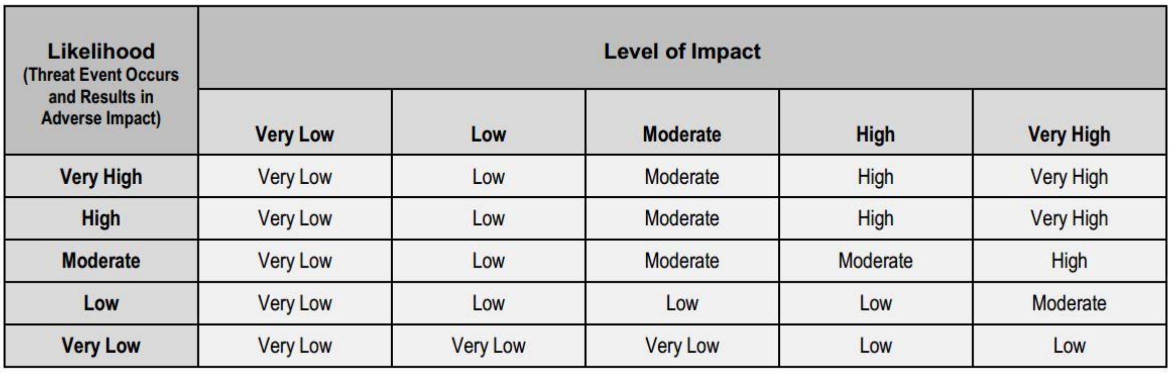

Threat Source
with Characteristics
(e.g., Capability, Intent, and Targeting for Adversarial Threats)
initiates
with Likelihood of Initiation
Threat Event
with Sequence of actions, activities, or scenarios
exploits
with Likelihood of Success
Vulnerability
with Severity
In the context of
Predisposing Conditions
with Pervasiveness
Security Controls
Planned / Implemented
with Effectiveness
causing
with Degree
Adverse Impact
with Risk as a combination of Impact and Likelihood
producing
ORGANIZATIONAL RISK
To organizational operations (mission, functions, image, reputation), organizational assets, individuals, other organizations, and the Nation.
Inputs from Risk Framing Step
(Risk Management Strategy or Approach)
Influencing and Potentially Modifying Key Risk Factors

Figure 7- Risk model, source: (NIST 800-30, 2011)

From the model we can extract that the risk is a function of the likelihood of a threat, vulnerability and potential impact:

Risk = Threat x Vulnerability x Impact (2.1)

This document provides several templates for risk assessments, however, without going too deep into the analysis of this text as it is out of scope of the present work, there is a relevant template that must be presented on this dissertation. It provides the means for assessing the level of risk by combining the likelihood and impact factors as presented in figure 8:



Likelihood (Threat Event Occurs and Results in Adverse Impact)	Level of Impact				
	Very Low	Low	Moderate	High	Very High
Very High	Very Low	Low	Moderate	High	Very High
High	Very Low	Low	Moderate	High	Very High
Moderate	Very Low	Low	Moderate	Moderate	High
Low	Very Low	Low	Low	Low	Moderate
Very Low	Very Low	Very Low	Very Low	Low	Low

Figure 8 - Assessment stage, level of Risk, likelihood and impact, source: (NIST 800-30, 2011)

2.7.2. International Organization for Standardization/International Electrotechnical Commission

ISO/IEC 27005 - Information Security Risk Management

The ISO/IEC[5] 27005, Information Security Risk Management, provides guidance into implementing risk management by supporting the requirements defined in ISO/IEC 27001 and unlike its counterpart ISO 31000 - which is a more holistic and general risk management framework - ISO/IEC 27005 is a specialized standard for managing risk in the information security domain and is a component of a continual monitoring and managing of risks and vulnerabilities that could create hazard to an organization.

This standard does not impose a particular approach to the implementation of a risk management system, but provides a detailed and iterative structure with six main processes to help tailoring an ISRM to an organization's needs as depicted in figure 9.

[5] ISO/IEC is an independent, non-governmental technical committee which belongs to the International Organization for Standardization and the International Electrotechnical Commission and it is based in Switzerland. Its purpose is to create international standards by ensuring strategic tools that increase productivity by promoting quality, safety and efficiency of products and services.

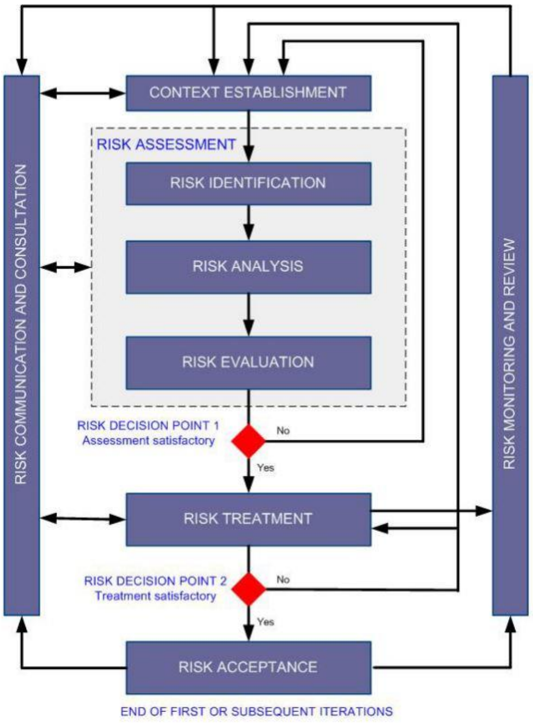


Figure 9 - Information security risk management process, source: (ISO/IEC, 2011)

The six main processes that characterize Information Security risk management are described as follow:

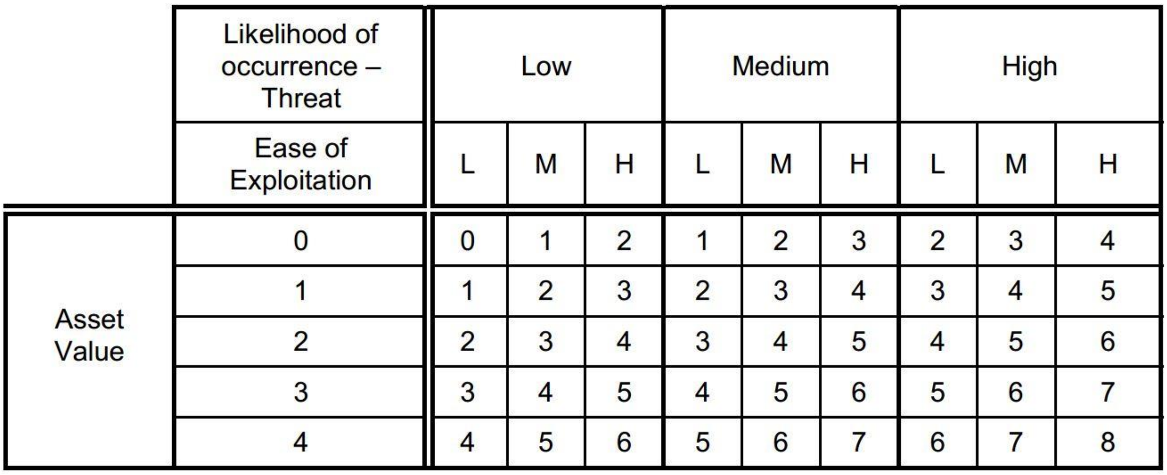
- **Context** – The first step of the iterative process begins by identifying the goals of an organization in context of information security needs and defining the scope and limitations of an information security management, such as risk evaluation, risk acceptance and impact criteria. This can be achieved by understanding the criticality of the assets involved, the law and regulatory requirements, stakeholders expectations, the cost, dimension and reputational damage imposed to the organization in case an information security incident happens (ISO/IEC, 2011);
- **Risk Assessment** is divided into three sub processes, namely the identification, quantification, qualification and prioritization of the risk evaluation as to meet the organization goals. At this step, the value of the organization's information assets is calculated, the business processes are assessed, the vulnerabilities and threats that exist or can exist are established, the IS controls are identified and is established a risk criterion where the identified risks are ranked and prioritized. The assessment step is comprised of three activities, the risk identification, risk analysis and risk evaluation.

At this stage, it is expected to identify all possible threats and vulnerabilities, exception handling, existing controls, assets, threat actors, threat landscape, consequences of losses of Confidentiality, Integrity and Availability (CIA), as well as the likelihood of a security event and the risk evaluation criteria. Risk assessment shall be performed periodically to identify changes on the threat landscape, vulnerabilities or organization's assets as to guarantee an up to date risk management.

- **Risk Treatment** is the process of modifying risk and is dependent on the outcomes of the risk assessment which provides four main non mutually exclusive risk treatment options which are risk modification, risk retention, risk avoidance and risk sharing (ISO/IEC, 2011).
    - **Risk modification** deals with the management to reduce or eliminate a threat or vulnerability by altering, removing or introducing controls.
    - **Risk retention** is based on the organization's policies and capacity to retain a risk, as such, if a level of risk falls into a risk acceptance criterion, the need to implement additional controls with associated costs is not necessary and the risk can be retained.
    - **Risk avoidance** is related with the benefits of treating a risk, if the cost or risk is to high that exceeds the benefits, the activities that potentially create the risk shall be avoided completely.
    - **Risk sharing** is the capability to share the risk of a certain activity with a third-party which is more capable of handling and manage that particular risk.
- **Risk Acceptance** should be handled at an organization's management level and treats the residual risk by assessing the outcomes provided by the risk opportunities, the risk modification cost or the risk effects.
- **Risk Communication** is based on strong and effective exchange of information, gathered on previous risk management activities, and agreed between major decision makers and different stakeholders so that an accepted outcome can be met in a coordinate way and in a continuous understanding environment.
- **Risk Monitoring** relates to the need of a constant risk monitoring and reviewing, form new added or changed assets in the management scope to new threats or vulnerabilities that enable continual alignment between business objectives, risk

acceptance criteria and risk management, therefore keeping the risk management relevant to the business goals.

The document provides several examples, guidelines and matrix templates to assess the information security risk in an organization. One of the relevant templates is the means by which an assessment can relate with the higher assets value in terms of actual costs of impact to an organization by the probability of a hazardous event. This assessment provides the effective mean to implement and direct a continuous monitoring on high ranking security controls. Figure 10 illustrates the matrix discussed:



	Likelihood of occurrence – Threat	Low			Medium			High		
	Ease of Exploitation	L	M	H	L	M	H	L	M	H
Asset Value	0	0	1	2	1	2	3	2	3	4
	1	1	2	3	2	3	4	3	4	5
	2	2	3	4	3	4	5	4	5	6
	3	3	4	5	4	5	6	5	6	7
	4	4	5	6	5	6	7	6	7	8

Figure 10 - Asset value vs vulnerabilities and threats levels, source: (ISO/IEC, 2011)

Chapter 3
Framework to Support Situational Awareness Dashboard

3.1. QlikView and Qlik Sense

The computational tool used in this work is the BI tool Qlikview, which allows to produce powerful applications with analytics and dashboards components. Qlik, former Qliktech, with its roots in Sweden, developed it back in 1993, though, has since moved to the United States of America.

Qlik defines it as a business discovery platform that provides self-service BI to users and enterprises(*What is QlikView?*) . Thus, Qlikview (and Qlik Sense, its sibling) is a data discovery, agile analytics and exploration tool (Howson et al., 2019) that allows consolidation of multiple and heterogeneous sources into tailor-made applications, which makes possible powerful strategic decision-making, based on a governed high performance associative engine. Qlik tools provides the means for a complete development of a BI tool, from extract, transform, load (ETL) of data to the development and designing of Dashboards.

QlikView offers much more in-depth configuration of objects in the front end and is usually considered as a guided business discovery tool in which the applications are usually developed by QlikView developers. Qlik Sense is more focused on business users and is considered as a self-service BI tool at the expense of a less customizable, but responsive, frontend. Both share the same engine running the data and both rely on a script for data modelling.

3.1.1. Qlik Associative Engine

Qlik tools differs from other BI tools on its core engine, which provides an associative experience, therefore not relying on predefined data paths or aggregations to navigate and explore data as OLAP cubes do (García & Harmsen, 2017). Usually, Structured Query Language (SQL) based technologies requires a level of data modelling which prompts to linear exploration and analysis of subsets of data, resulting in assumptions made in advance about what types of questions the users might have (Qlik, 2017). Contrary, Qlik associative technology allow the users to build complex analysis on the fly, without the need to go back at the data level to rebuild queries, providing interactive exploration and analysis, offering on-the-fly aggregation and calculation that enables the user to discover previously unforeseen insights which would have been missed by regular BI tools (Qlik, 2017).

This model provides analysis of complex relationships across the data model, because "every data point anywhere in the entire dataset to be analysed, regardless of how many data fields there are or how complex the underlying schema may be, should always be associated with all other data points at all times" (English, 2010).

Besides using this all natural to human thinking way of providing and associating data, Qlik uses in-memory data model, which stores all the application's data into primary memory, thus allowing faster response times and on the fly analysis (García & Harmsen, 2017) (see figure 11).

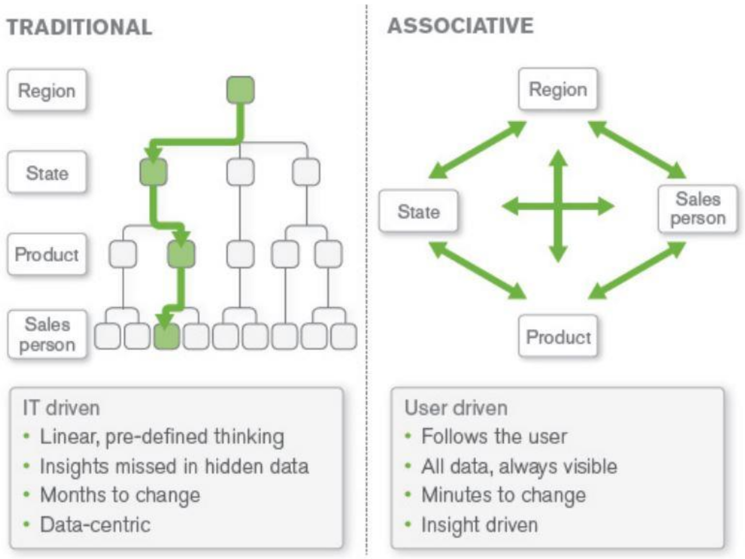


Figure 11 - Traditional BI solutions vs Qlik Associative model, source: (Qlik, 2014)

3.2. Situational Awareness Architecture

Cyber Situational Awareness shall provide comprehensive insights of the operational and business requirements of a given entity (Noel & Heinbockel, 2015) to a disparate array of users, providing organizational security status based on previously established metrics in automated processes for data collecting and information reporting (Dempsey et al., 2011) . Therefore, a Cyber Situational Awareness system is expected to be built on top of multiple data sources, thus a completely mixed and heterogeneous data input is anticipated. The information flow usually rises from diverse security-related assets, as defined by (Mell, Peter et al. 2012) and (Dempsey et al., 2011), continuous monitoring shall focus and be based on multiple resources, that is, on people, process, technology and environment assets, as illustrated in figure 12.

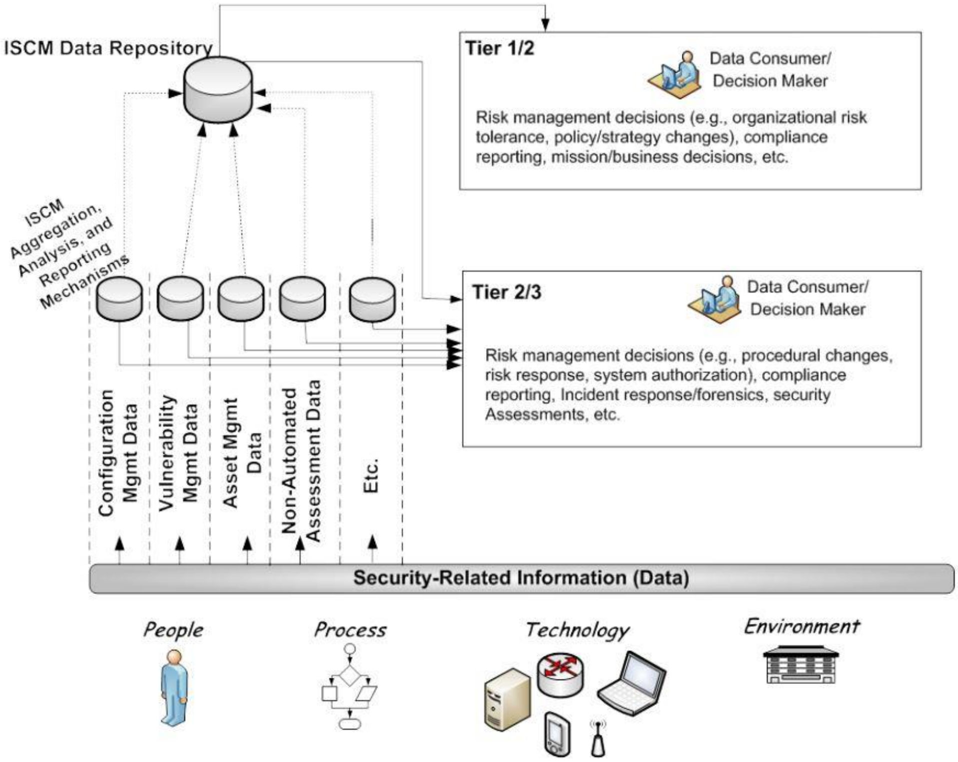

ISCM Data Repository
Tier 1/2
Data Consumer/ Decision Maker
Risk management decisions (e.g., organizational risk tolerance, policy/strategy changes), compliance reporting, mission/business decisions, etc.
ISCM Aggregation, Analysis, and Reporting Mechanisms
Tier 2/3
Data Consumer/ Decision Maker
Risk management decisions (e.g., procedural changes, risk response, system authorization), compliance reporting, Incident response/forensics, security Assessments, etc.
Configuration Mgmt Data
Vulnerability Mgmt Data
Asset Mgmt Data
Non-Automated Assessment Data
Etc.
Security-Related Information (Data)
People
Process
Technology
Environment

Figure 12 – ISCM, source: (Dempsey et al., 2011)

In addition, the target audience varies widely based on the needs and technical background of each user. A CEO will not need the same information that a CISO needs, nor will a security analyst.

Dempsey et al., (2011) defines a three-tier level organization wide approach to provide different types of information, interactions and risk management to and through the organizations personnel, be it at high-level security governance policy at tier 1 or at business processes in tier 2 and tier 3, which depicts information systems level information and processes. The Special Publication NIST 800-137 defines the three-tier model as follows (Dempsey et al., 2011):

- **Tier 1**, which is related with Organization, is where global risk management initiatives are conducted to define high-level information security governance based on the organization’s risk management, as it pertains to the organization as a whole. The activities on this tier includes defining how the organization access, responds and monitors risk management to ensure the effectiveness of the information security strategy. At this tier, security metrics and controls are expected to provide information regarding the decision making at a governance and strategic level.

- **Tier 2**, which is related with mission and business processes, defines how "core mission/business processes are prioritized with respect to the overall goals and objectives of the organization, the types of information needed to successfully execute the stated mission/business processes, and the organization-wide information security program strategy". Thus, business and security officials shall determine the critical assets and crown's jewels to monitor and secure, so that the business continuity is assured. Tier 1 and tier 2 make use of metrics and dashboards to assess, control and monitor information security controls from tier 3.
- **Tier 3**, the Information systems tier, is where the ISCM activities are related to a baseline level, i.e., the implementation and management of security controls at the technical and operational level are efficiently implemented and operating as needed, producing the desired results in accordance with the information security directives and requirements. At this level, it is expected to deal with security alerts, incidents and threat's reporting obtained from system-level controls.

When applying such implementation and strategy with an information security risk management framework, an overall and effective continuous monitoring can be achieved and provide an ongoing analysis of the security requirements and of each system contribution to the overall security state in a dynamic process that provides situational awareness for risk-based decisions.

With such an amount of data sources and information flow over the organization, and the need to provide tailor made information reporting in accordance with organizational objectives, policies and business rules, several SA models can be produced. A strategy to mitigate such a complex and time-consuming problem would be to identify reasonable and adequate security components and expand them as needed.

Thus, a classification and identification of critical business assets should be made from a top down perspective, it is of CEO and management 's responsibility to identify which assets loss, unavailability, disclosure or modification would impact most the business. From that point, an assessment between business management and information security management shall identify the criticality of these assets and the business impact taking in consideration the pillars of information, which means in terms of security, confidentiality, integrity, availability, non-repudiation and authentication.

Implementing and improving such a continuous monitoring tool would provide proactive information risk management to an organization, thus, allowing IT security users to monitor security controls more often and effectively, provide senior management employees with effective security reports and thus, actionable information, and ensure the opportunity in almost real time for risk-based decisions minimizing probable negative impacts.

3.2.1. Proof of concept architecture

As the final goal of this dissertation is to build a proof of concept based on a popular business intelligence tool that is capable of aggregating the data and providing visual data analysis, ensuring meaningful and useful information, a simple architecture and explanation is now presented with the intent to promote a simple solution to aggregate and present the data into a Situational Awareness dashboard.

The aim of the description is to provide a holistic overview, mapping the main IT components, processes and agents. It is intended that it shall have the capacity to support multidisciplinary sources, by providing an interoperability with other systems, with a high degree of reusability and extensibility, ensuring the use of transversal and consolidated data in the application that support the needs of information security monitoring with the ability to drill down from high level to aggregated information, with the appropriate and required granularity.

The situational awareness dashboard shall be based on automation, because automated solutions enhance efficiency, are less error-prone, therefore it is a more trustworthy way to monitor information security related information in a cost-effective way (Dempsey et al., 2011).

Figure 13, demonstrates the functional architecture idealized for this work using the Qlik tools. The first step is to create connections to multiple data sources located at stage 1. Qlikview and Qlik Sense contains multiple connectors to several types of data sources, from OLEDB and ODBC connectors to Web Services and Amazon Redshift, to name a few.

At stage 2, the Qlik architecture is presented which contains the ETL operations group followed by the presentation layer, where the dashboard is presented.

At stage 2.1, the data is extracted from the sources, whether it is data bases, text files, XLS files, or any other supported data type. The data can be extracted using regular SQL queries or, alternatively, using Qlik scripting language, which uses the same logic as SQL plus some native functions, names and syntax. As a best practice, the extracted data shall be saved in QVD files before any data transformation is applied. QVD files are Qlik's native format files that hold a table and are optimized for data compactness and speed of reading. These files can only be open on Qlik applications.

Saving the extracted data at this point would allow to have the raw data if any validation is required at the final stage of development.

Stage 2.2 is where the data is cleansed and transformed accordingly to the applications requirements. These data transformation operations should be performed using the Qlik scripting language as it is optimised to work on the application in-memory data. Usually, in Qlikview, best practices define that each table extracted and transformed should be saved in a separate file, a QVD file.

Stage 2.3 which defines the loading process, is where the data model is built on the previous data. At this step all previously transformed tables, which should be saved in QVD files, are loaded into a dimensional model. Dimensional models, such as star schema or snowflakes, represents an organization's business component and are composed of several dimension tables and, usually, one fact table. Star schema is the desired dimensional model for Qlikview as it allows optimized operations over the in-memory data.

Finally, at stage 4, which is the presentation layer, is where the visual objects are built and the dashboards are designed. These objects are composed of several expressions and dimensions and are developed in such way that allows for an exploratory data analysis. Qlik uses a native set language to define and develop the expressions called s*et analysis*. Set analysis allows to create expressions for dynamic exploration and calculation of data over the visual objects.

Notwithstanding, being out of the scope of this work, it must be noted that, Qlikview, is usually composed of 2 distinct layers. The Qlikview desktop, which is where a BI developer develops, designs and implement the Qlikview applications and the Qlikview Server. The Qlikview server is central on every QlikView architecture. It is at this stage that the applications are deployed, it where the applications are scheduled, be it for data loading or the distribution of the applications to the access points, where the end users can access applications. It is also at this stage that the user and application management, such as licensing and security, is made by a Qlikview administrator. For the purpose of this work, only a developer desktop version will be used, although the end product would be similar.

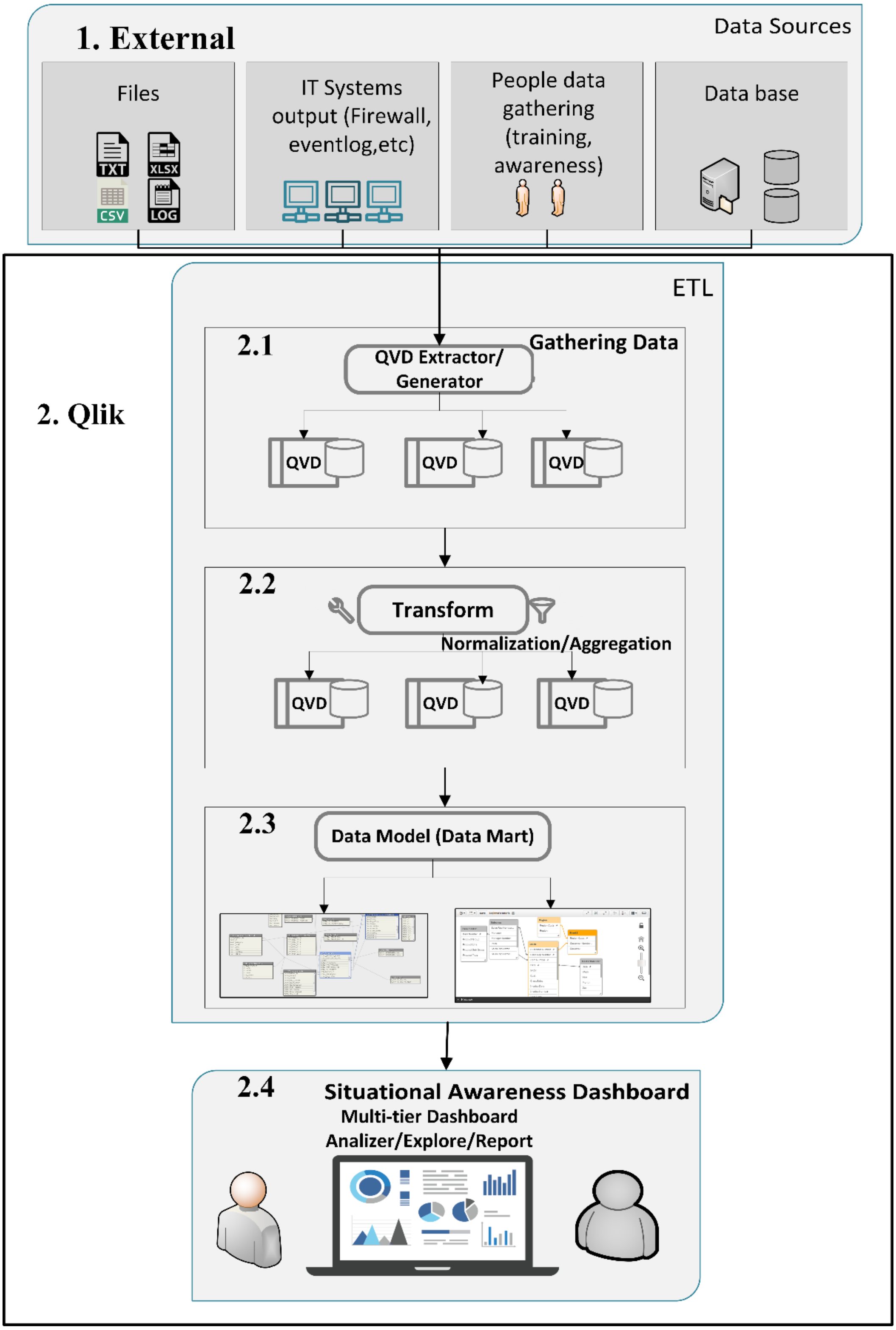


Figure 13 - Proposed functional architecture

Chapter 4
Information Security Metrics and Datasets

4.1. International Organization for Standardization/International Electrotechnical Commission 27001

In relation with information security, a series of standards were developed in the ISO 27000 family. ISO/IEC 27001 defines the foundations and requirements for establishing, implementing, deploying, monitoring, reviewing, maintaining, updating and improving an information security management system in which the full requirements to complete such system are listed, regardless of the size and nature of each organization and providing a tailor-made or customizable approach to such implementation based in an organizations' goals and business risk (ISO/IEC, 2018).

It provides the normative requirements for supporting the implementation and operation of an ISMS by defining a set of controls and security goals in which an organization can substantiate the development of an information security monitoring system by managing and reducing risk and complying to regulations, that contributes to confidence at providing an information security status of the organization therefore preserving the confidentiality, integrity and availability of information, based on each organization objectives, requirements, organizational processes and organizational dimension (ISO/IEC, 2013).

4.2. Security Metrics

The use of metrics provides overall insights about the state of affairs inside the business and/or the company. In addition, besides providing a snapshot of the past, it can also provide guidelines and helps shaping the pathway for the future as a way of supplying consistent information in a component of decision support systems. With the use of well-defined metrics, one can understand the characteristics and the nature of security controls or the implementation and monitoring of other security concepts, as risk management processes, incident management, vulnerability management, configuration management, security awareness and training, access control, firewall and other event logging, system monitoring, business continuity, physical security (ISO/IEC, 2016) , and thus providing value for the information gathered from all the data sources, for example, allowing an organization to understand if a certain risk has occurred and therefore mitigate potentially impacts on the business.

With today’s increasing number of data sources, massive amounts of data must be continuously monitored , but the risk of overlooking on the most important information, being error biased, and also, being extremely labour intensive, dictates more sophisticated and systematic methods of analysis, so that a qualitative outcome can be achieved in order to provide an evaluation and awareness of the situation in the organization, as such, a strong, simple and meaningful metric system shall be designed and implemented. With a correct set of security metrics, one can achieve an overall knowledge about the adequacy of the security controls and the protection of critical assets (Kott et al., 2014).

Using metrics that are properly designed and implemented provides monitoring, identification, evaluation, comparison and reporting of several security factors, providing decision taking with consistency, efficiency, objectivity in managing security risk in an organization in a continual improvement program (Barabanov et al., 2019). The use of metrics is essential to measure the success of a security awareness program, providing effectiveness and up to date contents to the SA program (PCI Security Standards Council, 2014). Providing effective security metrics and presenting them in a lean and objective visual way will enable an organization to demonstrate the value of an information security policy, allowing to show performance and improvements in information security domain, and providing accountability and compliance (Rathbun, 2009b).

International Organization for Standardization, (ISO/IEC, 2018) defines that the objectives of implementing an effective information security measurement are primarily:

- Evaluating the effectiveness of the implemented controls or groups of controls;
- Evaluating the effectiveness of the implemented ISMS;
- Verifying the extent to which identified security requirements have been met;
- Facilitating performance improvement of information security in terms of the organization’s overall business risks;
- Providing input for management review to facilitate ISMS-related decision-making and justify needed improvements of the implemented ISMS;

From a top-down approach, the needs of a security awareness application shall be identified in first place, which would guide towards the identification of the necessary metrics and associated measures in compliance with the organization needs.

International Organization for Standardization (Geneva, 2016), defines the process of monitoring, measure, analysis of data and consequent evaluation as a 6-step process:

- Identifying information needs;
- Creation and maintenance of measures (metrics);
- Establishing procedures;
- Monitoring and measurement;
- Analysis of results;
- Evaluation of information security performance and effectiveness;

A security metrics program shall be tailor made to a given organization, they must have content and context against a detailed situation and/or organization, aligned with the business and organizational goals. Determining the security level of an organization requires several steps to be done in order to achieve a consistent analysis (Kott et al., 2014): (1) What assets should be measured, (2) Organization of the requirements gathered on previous point and, (3) Construction of formulas (metrics) to answer about the security status.

To identify what assets shall be subjected to monitoring and, therefore, involved in the SA Dashboard, the value of the asset shall be calculated and the cost of the implementation shall be lower or, at least, have a high-Risk level to justify such an investment.

However, the quantification of each component might not be an easy and simple task and as (Kott, et al. 2014) noted that (Endsley, 1995) described a three step approach to calculate each element:

- Calculation of Asset Value: Based on quantifiable value of distinct information assets, often calculated by organizations, and depicted as the amount of spending during a time frame plus the depreciation/amortization value of the assets.
- Calculation of Potential Loss: Which are often related with five key types of breaches that might occur: confidentiality, integrity, availability, productivity and liability.
- Measurement of security spending: Organizations shall measure an enterprise-wide spending, albeit being extremely difficult, it can be split by business unit plus infrastructure.

In NIST 800-137, Dempsey et al., (2011) states that to determine metrics that evaluate and control ongoing risk in the organization, it shall "include all the security-related information from assessments and monitoring produced by automated tools and manual procedures, are organized into meaningful information to support decision making and reporting requirements. Metrics should be derived from specific objectives that will maintain or improve security posture. Metrics are developed for system-level data to make it meaningful in the context of mission/business or organizational risk management". It states that metrics shall be calculated from a variety of sources, from security monitoring, control assessment to security control data, obtained at different frequencies depending on the nature of the data sources. As examples, Dempsey et al., (2011) lists several metrics, as the number of unauthorized access attempts, contingency plan testing dates, number of vulnerabilities and severity of threats revealed or remediated and the number of users or employees who attended information security awareness training.

4.3. Metrics Concepts

For the development of the proof of concept, the goal of this work, several metrics will be selected across relevant literature. To construct meaningful metrics for a continuous monitoring systems, or situational awareness application, there will be a construction of a summary of metrics to answer some security controls or a direct implementation of metrics – for example in the ISO 27004 context – as such, an overview of ISO/IEC 27004 metrics, and Center for Internet Security, CIS Controls Measures and Metrics (Center for Internet Security, 2018) was evaluated and as a result a sample of some of these metrics will be presented and used in the POC.

The aim is to observe the most usual metrics, not only from a standard framework source as ISO/IEC 27000 but also from industry or academic approaches and to build a metrics summary that allows to have a foundation for the implementation of the proof of concept.

4.3.1. International Organization for Standardization/International Electrotechnical Commission 27004

The Standard ISO/IEC 27004 defines and provides the guidelines for the development and implementation of an information security measurement plan for assessing the effectiveness of an Information Security Management System implementation based on the specifications of ISO/IEC 27001. Thus, ISO/IEC 27004 defines and maps metrics for each control previously defined on ISO/IEC 27001. However, an organization can use the metrics defined even in the absence of such ISMS implementation, therefore it enables the possibility to measure controls other than defined on ISO/IEC 27001 by using the methodology provided by the standard (ISO/IEC, 2016).

Such methodology falls on the processes of enabling a set of activities or requirements such as how to identify the objects of measurement, what and when to measure and who shall measure, monitor and analyse such information (ISO/IEC, 2016).

4.3.2. NIST Special Publication 800-55 Revision 1

National Institute of Standards and Technology's SP 800-55 R1 supersedes NIST SP 800-55, Security Metrics Guide for Information Technology Systems and NIST Draft SP 800-80, Guide for Developing Performance Metrics for Information Security, and its aim is to define principles and guidelines to assist in the development, implementation and assessment of an efficient information security program. By efficient, it encompasses, through the use of measures and metrics, the adequacy of security controls and if they are in place, or instead, are non-productive, if information security spending is a justified investment, and prioritization of controls' implementation resulting in a cost effective and risk-based approach of an information security program. NIST 800-55 is based on the security controls established in NIST SP 800-53, and although it defines and are intrinsically connected with those security controls, it provides also the guidelines for a tailor-made definition and implementation of measures[6] to other security controls which an organization might implement. This document defines that such a program shall be based on four interdependent components:

- Strong Upper-Level Management Support – There is the need of commitment and support from top level management of an organization to ensure both budget for the program and to avoid organizational pressure. By providing such support, an information security program is a step further into a successful implementation.
- Practical Information Security Policies and Procedures ensure that the needed data is available to implement a measurement program. Without security policies, procedures and security authority it is impossible to have an information security management structure and thus, no data generation can be produced for such program.
- Quantifiable Performance Measures must be designed to provide meaningful insights about the security program performance. To reach such a degree of usefulness of information, the measures shall be designed in view of information

[6] NIST definition of measure is similar to the contemporary and more detailed definition of metric, therefore it will be used in this section interchangeably.

security goals of an organization, shall be readily and easily obtained, must be timely repeatable to provide trends over a timeframe so that it provides the performance tracking of information security controls.

- Result-oriented Measures analysis is related with the need to assess and analyse consistent data periodically transformed into meaningful information in a consistent way so that efficiency improvements and information security planning can be performed methodically.

In order to develop and implement a reliable information security measurement program, the Chew *et al.*, (2008)defines four critical success factors:

- Measures must yield quantifiable information (percentages, averages, and numbers);
- Data that supports the measures needs to be readily obtainable;
- Only repeatable information security processes should be considered for measurement;
- Measures must be useful for tracking performance and directing resources.

Notwithstanding that the type of measures that can be realistic obtained varies according to idiosyncrasies of an organization such as the information security program's maturity level and the implementation of security controls, the document focus on three types of measures:

- Implementation measures to measure execution of security policy – these measures are used to measure specific security controls, associated policies and procedures and system-level areas of interest. At first, the percentages measured by such metrics shall be below one hundred percent, but as the program matures, it is expected to reach 100 percent, therefore indicating that the information systems are in place and fully implemented, and the organization shall shift focus towards effectiveness/efficiency measures. Some examples of metrics in this context are: the percentage of information systems with approved system security plans; the percentage of servers within a system with a standard configuration; percentage of information systems with passwords policies configured as required.
- Effectiveness/efficiency measures to measure results of security services delivery – these metrics are used to measure the outcome of the implementation of security controls such as monitoring the implementation and operational aspects of a

program-level and system-level controls. Effectiveness is related with the robustness of the implementation whereas efficiency is directly related with the timeliness of the results. These metrics provide key performance indicators to decision makers about the policies and decisions, providing insights and continuous monitoring about the performance and effectiveness of an information security program. As an example of such metrics, the authors describe: the percentage of enterprise operating system vulnerabilities for which patches have been applied or have been mitigated, percentage of information security incidents caused by improperly configured access controls; percentage of system components that undergo maintenance on schedule.

- Impact measures to measure business or mission consequences of security events – these measures, besides measuring the impact of an information security program on an organization's mission, such as the cost savings produced by such program or costs incurred by security events, degree of trust perceived by the public, ultimately, evaluates the relationship between the investment on an information security program and the provisioning or available budget. As an example of a metric can be elaborated as the percentage of the organization's information system budget devoted to the information security program;

Acquiring and collecting relevant data to build measures and metrics is related with the maturity level of an organization's information security program, which in turn is defined by internal delineation of processes and procedures. As the program progresses and matures, the processes become more fine-tuned and standardized, thus providing more quality and quantity of data that can be used in the performance measurement

Finally, it is important to highlight in the context of the metrics thematic that this document defines the benefits of using measures as financial and organizational gains and improvements such as:

- Increase accountability: by helping to identify security controls that are inefficient, by this it means that, either they are incorrectly implemented, ineffective or not implemented at all.
- Improve information security effectiveness: an information security measurement program enables an organization to quantify and, therefore, justify information security investments into the implementation and improvements in securing information systems.

- Demonstrate compliance: Organizations can use the output of the measures with regulatory agencies, demonstrating evidence on keeping in compliance with laws, rules and regulations.
- Provide quantifiable inputs for resource allocation decisions: by allowing an organization in measuring the success or failures of past and current information security investments, it shall provide quantifiable data to support resource allocation and justify future investments allocated in accordance to a comprehensive risk management program.

4.3.3. Common Vulnerability Scoring System

Common Vulnerability Scoring System (CVSS) is a framework owned and managed by FIRST.Org, a north American non-profit organization which gathers a variety of computer security incident response teams from across industry, from governmental entities to educational and commercial organizations whose mission is to provide help to computer security incident teams around the world.

Common Vulnerability Scoring System, CVSS, is a standardized framework to rank the severity of computer systems security vulnerabilities, both in software, hardware and firmware. Being an industry standard, it ensures repeatable characteristics and helps the communication and scoring of vulnerabilities across organizations as it provides a platform agnostic vulnerability scoring methodology. CVSS computes the vulnerability, providing a quantitative value demonstrated by a numerical score, which foreshows the severity of a vulnerability relative to other vulnerabilities.

Hence, CVSS provides three benefits, it provides standardized vulnerability score, which leverages the management of a vulnerability, it provides an open framework, by defining consistent scoring towards vulnerabilities and, finally, it provides the enabling of prioritized risk-based actions, by creating computing environmental scores (FIRST, 2019a). It is essentially composed by three metrics groups: Base, Temporal and Environmental.

As explained in FIRST, (2019), the metrics groups are defined as follow:

- Base metric group represents the “intrinsic characteristics which are constant over time and assumes the reasonable worst-case impact across different deployed environments”;

- Temporal metric group is related with the representation of “changing and adjustment of the Base severity of a vulnerability that change over a period of time but not across user environments”;
- Environmental metric group represents the “characteristics of a vulnerability that are relevant and unique to a particular user’s environment”;

Base scores have the values assigned by a security analyst, then it is computed with a score assigned, ranging from 0 to 10, with 10 representing the most impacting and severe vulnerability. By having the base score defined, the temporal and environmental metrics can then be added as a weight part of the equation of the Base metric and so, it promotes refinement of the value of a vulnerability to a user’s environment (Environmental) and in a point in time (Temporal). The result of the computation of the metric equation is a vector string, a textual representation of the metric used to score such vulnerability and used to record and transfer CVSS metric information in a concise format. Some vulnerability metrics defined across other frameworks (ISO 27004, NIST) uses the CVSS scoring system to provide a quantitative value to vulnerabilities metrics.

4.3.4. Center for Internet Security

The framework was initially developed as the “Consensus Audit Guidelines” by the Center for Strategic and International Studies and tested successfully at NSA. It transitioned to SANS Institute as the “SANS Top 20 Critical Security Controls” and in 2015, it was assigned to the Center for Internet Security (Pescatore, 2017).

CIS critical security controls are the product of an effort made by the Centre for Internet Security consortium with the help from multiple information security experts from various sectors, including defence, government, healthcare, education, manufacturing and others. In short, by a community that (Controls, 2018) “share insight into attacks and attackers, identify root causes, and translate that into classes of defensive action. Document stories of adoption and share tools to solve problems. A community that tracks the evolution of threats, the capabilities of adversaries, and current vectors of intrusions. That map the CIS Controls to regulatory and compliance frameworks and bring collective priority and focus to them. A community that share tools, working aids, translations and identify common problems (like initial assessment and implementation roadmaps) and solve them as a community.”

Given the multitude and disparity of security tools, technologies and standards available today, the CIS controls were built upon the necessity of summarizing and focusing on the most fundamental and valuable actions that every organization shall take in order to prevent, alert and respond to the attacks that organizations fall victims these days, therefore they representing a set of security actions that are focused and prioritized in accordance with industry and government security requirements.

CIS controls are built around five main principles:

- Offense informs defence: in a lesson learn style, the knowledge base from attacks shall be used as effective source for building practical defences.
- Prioritization: the focus shall be on implementing controls that provide the biggest risk reduction in the simplest way against the biggest threat actors.
- Measurements and Metrics: a set of common metrics shall be implemented in order to communicate the security status across the distinct type of an organization's personnel, from IT specialist to executives.
- Continuous diagnostics and mitigation: an effort shall be made on continuous monitoring the implementation and validation of the information security status in order to guide the organization towards the security goals.
- Automation: will drive the organization efforts on the correct implementation to the security controls.

CIS security controls, albeit having a well-defined set of security actions, shall be adapted to an organization's needs and have the implementation in a step by step method, depending on the needs and realities of an organization.

There are twenty CIS controls, which are divided amongst three main groups: Basic, Foundational and Organizational (Controls, 2018):

- Basic
  - Inventory and control of hardware assets;
  - Inventory and control of software assets;
  - Continuous vulnerability management;
  - Controlled use of administrative privileges;
  - Secure configuration for hardware and software on mobile devices, laptops, workstations and servers;
  - Maintenance, monitoring and analysis of audit logs;
- Foundational

  - Email and web browser protections;
  - Malware defences;
  - Limitation and control of network ports, protocols and services;
  - Data recovery capabilities;
  - Secure configuration for network devices, such as firewalls, routers and switches;
  - Boundary defence;
  - Data protection;
  - Controlled access based on the “need to know”;
  - Wireless access control
  - Account monitoring and control
- Organizational
  - Implement a security awareness and training program;
  - Application software security;
  - Incident response and management;
  - Penetration testes and red team exercises;

4.4. Meaningful Metrics for Cyber Situational Awareness proof of concept

As stated on the above section, relevant metrics were gathered from well-established frameworks, with a focus on simplicity and future improvements. Since there are quite some limitations on gathering free and available information security data, not all metrics can be implemented in the POC and, when possible, some sources will be constructed with customized data based on the required elements for a given metric. In appendix A, a compendium of all collected metrics from relevant authorities is directly extracted from the sources and summarized in a table. For simplicity, the implementation of the dashboard will have the focus on metrics that can be accomplish from data accessed in an automated way, therefore, data from written tests, assessments or surveys won’t be considered into this work. This type of data needs a middle layer implementation, from a *human interface* to computer data - a digitalization process - as for example insertion of data from written surveys into digital content, which is out of scope of the subject but should be taken into consideration for future developments or evolution of the dashboard.

As previously stated, heterogeneous data sources are expected, and the implementation design will be such that any kind of (previously) processed data can be inserted into the data model of the SA Dashboard through the ETL process in a tailor-made method. Therefore, this type of data can and would be considerate in any real-life application, if the organization understand that there is value on the identifying and control of such information. On the following subsections, several metrics will be presented and explained with the main objective to integrate in the POC and provide a simple, yet meaningful dashboard with actionable information that shall serve as a starting point for a robust dashboard. Each metric definition will be expressed in terms of mathematical symbols to avoid ambiguity and to maintain a pattern across all presented metrics in the subsequent sections.

4.4.1. Training and Awareness

In the People - Process - Technology triad concept, people are often considered the weakest link in information security. Several factors contribute to this attribute, be it lack of training and knowledge in the cyber security domain or simply negligence for security practices.

Security shall start with end users, as the level of security awareness is a definitive source for the organization's security as a whole (Korpela, 2015). Therefore, several measures can be taken to reduce the risk of human factor, as for example, providing mandatory training on security and awareness to all employees, including C-Suite, conducting awareness tests, for example, periodically developing an organization's internal phishing simulation test to measure the number of users that open the email and follow the links, how many users reported the email, number of stolen or lost devices reported before and after training deadline.

More often than not, employees are not aware of the cyber risk involved in common operations and of their daily life working procedures, and the potential reputational, economical and legal repercussion that an organization might come into for a misleading action. Plus, cyber-attacks and social engineering are becoming more prevalent in an increasingly interconnected world. Furthermore, employees not only should be trained to identify potential incidents, but also shall be trained in how to respond to a possible cyber-attack, and therefore, be able to mitigate the risk of loss or avoid disruption. A user shall be trained to be aware and to identify potential or of typical types of threats, for example, ransomware, malware, phishing campaigns, digital identity theft or recognizing social engineering tactics.

A strong training in security awareness, which shall establish a minimum cyber security awareness and provide an up-to-date set of policies in a continuum evaluation and evolution of those rules and policies, followed by periodic assessments on human vulnerability is of utmost importance. However, with this investment comes several questions as how to measure the deployment of an awareness program, the return on investment, or if the human behaviour is being changed by the spending on training, these questions can be translated into actionable metrics and help providing return of investment, both monetarily and security-related efficiency:

- Have the employees completed the training sessions?
- What is the role/position of employees that are not complete the training?
- Is our staff prone to social engineering *scams*?
- What is the number of incidents reported after vs before awareness training?

Human Factors Metrics

This base measure provides the total number of users that finished the training assessment until the deadline and an organizational global overview of the target users that were intended to assess a security awareness training, it is defined as follow:

1. What is the total numbers of users completing the training?

$$UC = \sum_{i=1}^{n} x_i \tag{4.1}$$

Where

UC denotes users completing training;

x is the number of users completing the training sessions;

From this measure, several metrics can be achieved which will allow to answer further questions and apply corrections on course directions and training strategies

2. What is the percentage of users completing the training?

$$\%UCT = 100 \times \frac{\sum_{i=1}^{n} y_i}{\sum_{j=1}^{m} t_t} \tag{4.2}$$

Where:

%UCT represents the percentage (%) of employees completing training;

y is the number of users completing the training sessions;

t is the number of number of users;

This metric can further be analysed by several dimensions, as for example, by business unit or department, by region, and so on. This metric provides useful insight and can give an overview about training behaviours and allow for further corrections into the training program.

3. What is the percentage of users completing the training in Year over Year comparison?

$$\%UCTY = 100 \times \frac{\sum_{i=1}^{n} y_i}{\sum_{j=1}^{m} t_t} \tag{4.3}$$

Same as metric 4.2, but over Year dimension

This metric provides a general idea about the implementation of the security training program, the adherence to, and evolution of the aforementioned program. A good indicator should be when an increase of the completion of the training is seen in most recent years.

4. What type of users don’t complete the training, by staff dimension:

$$\#UNC = \sum_{i=1}^{n} w_i \tag{4.4}$$

Where:

UNC represents the total number of users not completing training;

w is the employee not completing the training session;

This metric is calculated over Department/Business Unit dimension and/or by employee role and provides visibility over the of users lacking on training schedules. All users shall be aware and subjected to information security topics, starting from top management, which are the most accountable for the security of the information assets.

5. What is the percentage of users not completing the training?

$$\%UNCT = 100 \times \frac{\sum_{i=1}^{n} w_i}{\sum_{j=1}^{m} t_j} \quad (4.5)$$

Where:

UNCT indicates the percentage (%) of users not completing training session;

w is the number of users not completing training;

t is the number of users;

6. Percentage of Users completing the training and passing the test.

$$\%UPT = 100 \times \frac{\sum_{j=1}^{m} f_j}{\sum_{i=1}^{n} x_i} \quad (4.6)$$

Where:

UPT represent the percentage (%) of users passing training test;

f is the number of users passing the training test;

x indicates the number of users that made the training;

This metric can be analysed over several dimensions and provides a fast and aggregated proportional overview of training completion.

7. Complex password:

$$\%UQP = 100 \times \frac{\sum_{i=1}^{n} c_i}{\sum_{j=1}^{m} p_j} \quad (4.7)$$

Where:

%UQP is the percentage (%) of complex passwords;

c represents complex passwords;

p is the number of passwords;

The goal of this metric is to obtain information about the ratio of passwords that meets the organization policy. However, being related with security awareness training, it is also related with the enforcement, or the need to enforce, of organizational security policies.

8. Quality of passwords, *crackable*:

$$\%UNCP = 100 \times \frac{\sum_{i=1}^{n} c_i}{\sum_{j=1}^{m} p_j} \tag{4.8}$$

Where:

%UNCP is the percentage of non *crackable* passwords;

c is the number of non *crackable* passwords;

p represents the number of passwords;

This metric provides a ratio of the current *crackable* passwords in a temporal window, ISO/IEC 27004:2016 (2016) recommends a window time of 4 hours to crack and actions to be taken if the ratio is below 0.8.

Social Engineering Metrics

The data feed for these types of metrics shall be gathered from the IT service desk, which shall keep track of the incidents reported and categorize them accordingly, another source of data is also the output of periodic social engineering tests.

For example, information security units can forge fake emails with typical social engineering tactics, like phishing emails, and monitor the number of users that *click*/follow the email links. The data collected shall be stored and can be used to extract relevant information. Following a security training, it is expected to have a bigger percentage of awareness – minor number of users following phishing email - from the users after these training sessions, but it is also expected to have the number lowering and became stabilized along the time. As such, this indicator can provide guidelines for the importance of these type of trainings and the frequency needed to assess such training sessions to the organization users.

9. Percentage of users following a phishing email.

$$\%UPH = 100 \times \frac{\sum_{j=1}^{m} f_j}{\sum_{i=1}^{n} u_i} \tag{4.9}$$

Where:

%UPH represents the percentage (%) of users following a phishing email;

f is the number of users following the link on phishing email;

u is the number of users that were subjected to the fake phishing email;

This metric shows the proportion of users that are unaware of common social engineering tactics and follow/click the link on a fake test phishing email. It provides information regarding the efficiency of the security awareness training.

10. Percentage of reported phishing emails.

$$\%URPH = 100 \times \frac{\sum_{j=1}^{m} r_j}{\sum_{i=1}^{n} u_i} \tag{4.10}$$

Where:

%URPH denotes the percentage (%) of reported phishing email;

r is the number of users reported phishing email;

u represents the number of users that were subjected to the fake phishing email, which corresponds to the total number of emails sent;

Similar to metric 1., and again presents information about the efficiency of the security awareness training and how users are more prone to scam tactics and to report such incidents.

11. Number of reported incidents related with social engineering topics.

$$USEI = \sum_{i=1}^{n} s_i \tag{4.11}$$

Where:

USEI is the total number of social engineering incidents;

s represents social engineering incidents;

An increase in this metrics can be a valuable indicator of the awareness of the users about social engineering topics, but also can indicate that the organization is getting more exposed to this type of attack, to overcome this situation, a correlation can be done with other incident metrics.

Table 1 summarizes the training, awareness and social engineering metrics, regarding the sensor group type to which they belong.

Sensor Group	Definition	Metric Definition
Training and awareness	# users completing training sessions	# = ∑ (users completing training)
Training and awareness	% of users completing training sessions	% = (∑ users completing training / ∑ total users) x 100
Training and awareness	% of users completing the training in Year over Year	% = (∑ users completing training / ∑ total users) x 100
Training and awareness	# of users not completing training sessions over dimension	# = ∑ users not completing training
Training and awareness	% of users not completing training sessions	% = (∑ users not completed)/(∑ total users) x 100
Training and awareness	% of users completing the training and passing the test	% = (∑ users passing)/(∑ total completing training) x 100
Training and awareness	% Complex passwords	% = (∑ Total C=2)/(∑ Total C= {1;2}) x100
Training and awareness	Quality of passwords	% = (∑ non crackable passwords)/∑(Total passwords) x100
Social Engineering Metrics	% of users following a phishing email	% = (∑users following link)/(∑users subjected to test email)x100
Social Engineering Metrics	% of reported phishing emails	% = (∑reported emails)/(∑users subjected to test email)x100
Social Engineering Metrics	# of reported incidents related with social engineering	# = ∑ Social engineering incidents

Table 1 - Training, awareness and social engineering metrics

4.4.2. Vulnerability

This metrics group provide information regarding the information security landscape´s vulnerability to malicious attacks (ISO/IEC, 2016), identifying known vulnerabilities and the scan coverage of assets, therefore permitting the remediation of such vulnerabilities and minimizing the opportunity window for attackers (Controls, 2018).

12. Percentage of critical systems where vulnerability and *pentest* assessment has been done after major releases.

$$\%UPVA = 100 \times \frac{\sum_{j=1}^{m} a_j}{\sum_{i=1}^{n} c_i} \tag{4.12}$$

Where:

%UPVA is the percentage (%) of pentest/vulnerabilities assessments;

a represents the number of pentest or assessment performed over IS;

c is the number of information systems classified as critical;

The information provided by this metric allow the control over the assets classified as critical and the possible vulnerabilities to malicious attacks, therefore allowing the remediation plans to be implemented.

13. Unpatched vulnerabilities.

$$\%UV = 100 \times \frac{\sum_{j=1}^{m} u_j}{\mathrm{t}} \tag{4.13}$$

Where:

UV represents the percentage of vulnerability unpatched;

u is the number of unpatched systems;

t is the total number of systems;

General overview of the percentage of unpatched vulnerabilities.

14. Ratio of organizational vulnerabilities coverage.

$$ROVC = \frac{\sum_{j=1}^{m} a_j}{\sum_{i=1}^{n} s_i} \tag{4.14}$$

Where:

ROVC represents the ratio of vulnerability assessments performed on assets;

a is the number of system where vulnerability assessments were performed;

s is the number of systems;

This metrics provides Evaluation of the organizational visibility over vulnerabilities landscape.

4.4.3. Incident Handling

The information provided by this metric's group allows to effectively manage the organization's reputation and information assets by controlling an incident response implementation, including events and weaknesses (ISO/IEC, 2013).

15. Cost related with lack of Information Security

$$SCSI = \sum_{i=1}^{n} c_{i} \tag{4.15}$$

Where:

SCSI represents the sum of costs of information security incidents;

c is the cost of each security incident during time frame;

This metric gives awareness over the total cost of lack of information security, it is useful to quantify ROI over the implementation of security controls and related SA.

16. Actions for security improvements

$$NSI = \frac{\sum_{j=1}^{m} a_j}{\sum_{i=1}^{n} l_{\mathrm{i}}} \tag{4.16}$$

Where:

NSI represents the number of security improvements after SI incidents;

a = Number of security incidents that trigger actions;

l = Number of security incidents;

Overview on whether security incidents triggers security improvements.

17. Effectiveness of incident management

$$TINS = \sum_{i=1}^{n} u_i \tag{4.17}$$

Where:

TINS represent the total number of SI incidents not solved over an agreed time frame by category;

u is the number of incidents not solved on agreed timeframe;

This metric measures the effectiveness of the implementation of security incident response, thus allowing adjustments on the ongoing approach to incident response.

18. Security incidents

$$TISI = \sum_{i=1}^{n} l_i \tag{4.18}$$

Where:

TISI it the total number of security incidents over a time frame by category;

l represents the number of security incidents over a timeframe;

This metric provides information over the total number of information security incidents over a period. This metric should be constructed on an agreed timeframe, over category dimension to provide an in-depth analysis of the trends of incidents.

4.4.4. Assets Inventory

This metric group is responsible for retrieving information regarding an organization's information assets, especially the crown's jewels of such an organization.

The aim is to identify possible unauthorized devices connected to the network, identify the authorized ones, and most important, identify and keep track of critical business assets, those which the disclosure, modification or unavailability could create impact on the organization. The process of identification and classification should be taken in a joint effort between information security and business teams as to ensure the confidentiality, integrity and availability of such assets.

19. Authorized devices

$$TAD = \sum_{i=1}^{n} a_i \tag{4.19}$$

Where:

Tad represents the total number of authorized devices;

a denotes authorized devices;

This measure provides the total number of authorized devices in the inventory. It shall be accompanied by a level of detail in which the identification, IP address and security status is shown.

20. Unauthorized devices

$$TNAD = \sum_{i=1}^{n} u_{\,i} \tag{4.20}$$

Where:

TNAD indicates the total number of unauthorized devices;

u denotes unauthorized devices;

This metric provides the total number of unauthorized devices in found in an active discovery tool and provides the information need so that either the device is removed or the inventory is updated. To accomplish the goal of this metric, the details, similar to those of the authorized assets, shall be shown in tabular form to provide the relevant information to proceed with the required actions, be it an update or removal action.

21. Critical assets

$$TCA = \sum_{i=1}^{n} c_i \tag{4.21}$$

Where:

TCA specifies the total number of critical assets;

c is the critical asset;

Provides the total number of critical assets and along with appropriate level of detail, provides information regarding multiple specificities of the asset. The details should include the identification, IP address, risk level, protection concept and other relevant information.

22. Percentage of critical assets protection/patch

$$\%PCAP = 100 \times \frac{\sum_{j=1}^{m} p_j}{\sum_{i=1}^{n} c_i} \tag{4.22}$$

Where:

%PCAP denotes the percentage (%) of critical assets in conformity with protection/patch concept;

p is the number of critical assets with protection;

c is the number of critical assets;

Table 2 summarizes all the previous defined metrics. These metrics belong to the vulnerability, incident handling and assets inventory group.

Sensor Group	Definition	Metric Definition
Vulnerability	% of Critical sys. vulnerability and pentest assessment	`% = ∑(pentest done on CIS/∑Critical information systems)x100`
Vulnerability	Unpatched vulnerabilities	`Score = CVSS Score x Affected System`
Vulnerability	Ratio of organizational vulnerabilities coverage	`S:AS = ∑( Systems Vuln. Assessemnt./∑Systems)`
Incident Handling	Cost of lack of information security	`#$ = ∑(Costs IS incidents)`
Incident Handling	Security improvements	`Simp:Sinc = ∑(Sinc trigger actions)/∑(Sec incidents)`
Incident Handling	Effectiveness of incident management	`# = ∑(Incidents not solved)`
Incident Handling	Security Incidents	`# = ∑(Security incidents)`
Assets Inventory	Authorized devices	`# = ∑(Authorized devices)`
Assets Inventory	Unauthorized devices	`# = ∑(Unauthorized devices)`
Assets Inventory	Critical Assets	`# = ∑(Critical assets)`
Assets Inventory	% Protection/patching	`% = ∑(Critical assets protected)/∑(Critical assets)x100`

Table 2 - Vulnerability, Incident and Inventory metrics

4.5. Security Index

This section presents a mathematical effort to quantify the level of a security program implementation. Any organization that implements a security program deals with an enormous amount of distinct security metrics in different formats, being absolute numbers, statistical, percentages, over distinct partitions as region or business units, and as such, a common index shall be seek to provide a quantifiable score (Pareek, 2017).

The security index is a metric defined based on all security metrics that provides a high-level assessment of the situational awareness on ongoing basis. It is defined as a one number only on a scale from zero to one, with the purpose to provide a global overview of the state of Security implementation. This type of KPI is rather useful in a context of C-suite presentation because it provides a unique index.

$$\mathrm{SI} = \sum_{i=0}^{n} w \times \mathcal{M}_i \tag{4.5.1}$$

Where SI is the Security Index, n is the total number of metrics evaluated, w is the attributed Weight and $\boldsymbol{\mathcal{M}}$ is the evaluated metric. Due to the disparity of metrics computed and the absence of threshold defined for each one, Z Scores calculation of each metric will be used in order to obtain standardized scores (Pareek, 2017). As such, $\boldsymbol{\mathcal{M}}$ is extended as:

$$\mathrm{Z_i} = \frac{x_i - \mu}{\sigma} \tag{4.5.2}$$

The perceived weight is the risk score attributed to each metric in the organization context and shall be aligned with business goals and risk management. Albeit several definitions for risk scoring in information security are presented, for the sole purpose of simplicity in this work it is assumed that risk score in based on the Risk equation (1)[7] which states that the Risk level is a function of threats, vulnerabilities and impact (Kott et al., 2014).

4.6. Dataset construction

This section aims at documenting the creation of the datasets as well as the rules and formulae used to populate the fields.

Information security deals with core information regarding organizations, therefore the open access to real data is usually difficult. This data might expose organization's

[7] Previously defined in section 2.8.3

vulnerabilities or policies that should not be disclosed to individuals outside them. Therefore, rather than using third party datasets, e.g., the ones available at academic repositories, such as “Impact Cybertrust” or “Awesome-Cybersecurity-Datasets”, just to mention a few, several customized datasets were assembled to populate the final application. There was the need to create files to simulate data extracted from a data base or other operational source systems. The intention to create random data in order to avoid creating bias was taken into account and is explained in each subsection. Hence, randomness was introduced to provide a more realistic approach. The synthetic datasets produced in this thesis will be made available for the community such that, at a later time, can be used for development and testing of new algorithms and processes.

An hypothetic organization was idealized, with its offices distributed in four countries, composed of 2007 employees (users) distributed across different locations and with a total of presented 6553 devices, representing from laptops to other hardware connected to the corporate network. As previously explained, due to the lack of industry or corporate datasets to implement the product of this work, several metrics were selected in accordance to the possibilities of creating the customized data to work on the selection of metrics. Metrics is a hard-working process subjected to several assessments, discussions, advances and setbacks, and is always done in accordance to the organization’s goals and vision

The first dataset, which is pertaining to Training and Awareness data, is explained in a more in-depth method to elucidate the steps followed on the creation of the data. The following datasets will be explained in a more general context as the creation method is similar to the first and the aim is to avoid repetitions in the explanation.

4.6.1. Training and Awareness Dataset Construction

The aim of this dataset is to provide data to analyse in accordance to the metrics defined in subsection 4.2.1. A total of ten fields were projected:

- **RowID**: which is a sequential ID for the table;
- **UserID**: a unique random user identifier;
- **Country**: corresponds to the workplace/office country of the employee;
- **Date**: which is the date of the assessment and the files will be generated for 2 semesters and 2 years.
- **Mandatory**: populated with 0 or 1, 0 for not mandatory, 1 for mandatory;

- **Completed**: this fields identifies if the user/employee completed the training;
- **Score**: The score on training examination, if the value is below 60, the user is considered to have failed the test.
- **C Suite**: value 0 or 1, when the value is 1, it indicates that the user belongs to C Suite Category;
- **Mngmnt**: Similar to the above description, but 1 indicates that the employee is of management category;
- **Employee**: Similar description as above, when value is 1, the employee if of general employee category;

4.6.2. Training and Awareness Dataset Fields Rules

This section presents the general rules and associated formulas used to create the customized data:

- Field **RowID**, sequential number (for 2007 users);
- Field **userID**, generated for 2007 users:

  User id must have a length of 10 characters composed of both numbers and letters, obeying to the following rules:

  - The alphabet will be defined as `∑={C,D,E,F}` and over the following discrete number set S `= [0,..,9]` for all options henceforth;
  - The first position of the user ID is always a character;
  - The second position is a randomized iteration over either a character or a number;
  - The third, fourth and fifth positions in the user ID are always a random generated number;
  - The second position is randomized iteration over either a character or a number;
  - The seventh, eighth, ninth and tenth position is always a random generated number;
- **Country**: Portugal, France, Spain, Brazil

  Random selection of 52% of user's population for Portugal:

  - First, a sequential row ID is generated for all population (o to 2006);
  - A random number is generated for all row id on an adjacent column using the following formula:

    ```
    +RAND()
    ```

  - Lastly, 52% of the population is selected by using:
    ```
    =+IF(ROW()>52%*COUNTA($A$2:$A$2007);"";INDEX($A$2:$A$2007;RANK.AVG(H2;$H$2:$H$2007;0);1))
    ```
  - After this process is concluded, a VLOOKUP: is applied on the original table by using row id and the selected sample of users is removed from the sample generating the population:
    ```
    =VLOOKUP($A$2:$A$2007;Sheet3!$A$1:$A$1042;1;FALSE)
    ```
  - To select Spain and France samples, the algorithm is used on the other candidates, on the remaining population that did not got selected on the previous iteration, repeating from step 1 to step 4, with the difference on the percentage value for each country:
    - Spain:
      ```
      =+IF(ROW()>53%*COUNTA($A$2:$A$966);"";INDEX($A$2:$A$966;RANK.AVG(G2;$G$2:$G$966;0);1))
      ```
    - France:
      ```
      =+IF(ROW()>18%*COUNTA($A$2:$A$2007);"";INDEX($A$2:$A$2007;RANK.AVG(H2;$H$2:$H$2007;0);1))
      ```
    - Brazil will be populated with the remaining users;

- **Date**: This field represents the date of assessment. It was randomly generated between dates 01.01.2018 and 31.01.2018 with the following formula:
  ```
  =+RANDBETWEEN(DATE(2018,1,1),DATE(2018,1,31))
  ```
  - If user did not complete the training, then the end of the month is introduced, `31-01-2018.`
- **Mandatory**: the value of this field is always 1. This means that the training is mandatory for all type of users.
- **Completed**: This field flags if a user did complete a security training (`1`) or, on the contrary, the user did not complete the training (`0`). It was randomly calculated with a value of 10% of the users not completing the training, using the following formula:
  ```
  =+IF(ROW()>10%*COUNTA($A$2:$A$2007),"",INDEX($A$2:$A$2007,RANK.AVG(B180,$B$2:$B$2007,0),1))
  ```
- **Score**: this field presents the score a user obtained in the Training test.
  - If a user has a value of zero on the Completed field, then it is populated with a `0` value.

  - For all the other cases, Normal Distribution was used to calculate the score on the training test. An optimistic mean of 75 was used with a standard deviation of 10, with the following Excel formula:

    ```
    =+NORMINV(RAND(),75,10)
    ```

- Fields type of **Employee Category**:
  - C Suite, Mgmt. and Employee were randomly calculated using the following formulas respectively:
  - For C suite employees, a 1% was used:

    ```
    =+IF(ROW()>1%*COUNTA($A$2:$A$2007),"",INDEX($A$2:$A$2007,RANK.AVG(B2,$B$2:$B$2007,0),1))
    ```

  - For Management employees, field “Mngmnt”, a 1% was used, excluding the previous selected users:

    ```
    =+IF(ROW()>9%*COUNTA($A$2:$A$2007),"",INDEX($A$2:$A$2007,RANK.AVG(B180,$B$2:$B$2007,0),1))
    ```

  - (regular) Employees were populated using all the users that did not populate in the preceding samples.

A total of 4 files were created, simulating the results for Information security training in a total of 2 years, 2018 and 2019, and 2 semesters. All the scores were generated using Normal Distribution. The field ‘completed’ was randomly calculated also, but for 2018 second semester it was used 8% of the population, for year 2019, first semester a value of 7% and on the second semester, a value of 3% of the population was attributed the value 0 (not completed). In figure 14, the results of the distribution by country generated in the dataset is presented. Table 3 summarizes the rules and associated formulas used.

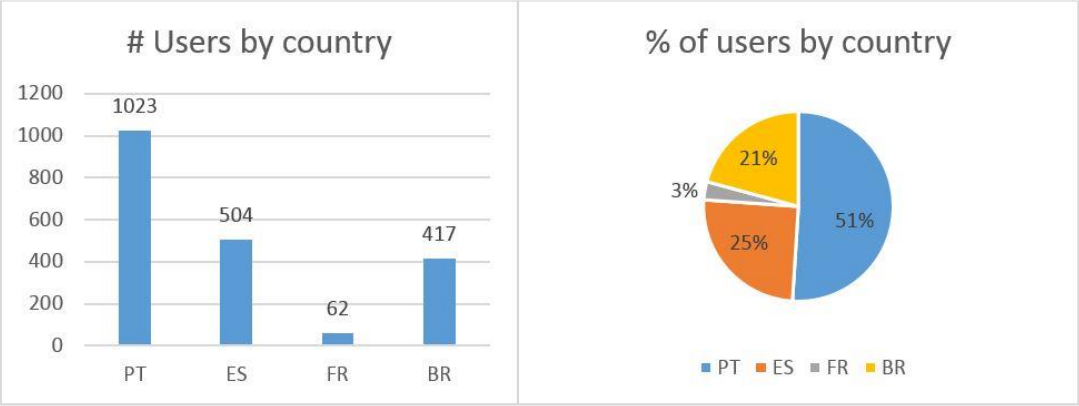


Figure 14 – Views of the users distribution by country that was assumed for the dataset

Rules	Type	Position	Formula
∃x ⊂ {C;D;E;F}	Char	1	CHAR(RANDBETWEEN(67;70))

∃x ⊂ {C;D;E;F}Y ∃x ⊂ [0,9]	Char
	||
	Int
		2	CHOOSE(
			(RANDBETWEEN(1;2));
			CHAR(RANDBETWEEN(67;70));RANDBETWEEN(0;9)
			)

∃x ⊂ [0,9]	Int	3	RANDBETWEEN(0;9)
∃x ⊂ [0,9]	Int	4	RANDBETWEEN(0;9)
∃x ⊂ [0,9]	Int	5	RANDBETWEEN(0;9)

∃x ⊂ {C;D;E;F}Y ∃x ⊂ [0,9]	Char
	||
	Int
		6	CHOOSE(
			(RANDBETWEEN(1;2));
			CHAR(RANDBETWEEN(67;70));RANDBETWEEN(0;9)
			)

∃x ⊂ [0,9]	Int	7	RANDBETWEEN(0;9)
∃x ⊂ [0,9]	Int	8	RANDBETWEEN(0;9)
∃x ⊂ [0,9]	Int	9	RANDBETWEEN(0;9)
∃x ⊂ [0,9]	Int	10	RANDBETWEEN(0;9)

Table 3 – Summary of User ID generation rules

4.6.3. Social Engineering Dataset Construction

This dataset file provides data to create information regarding social engineering testing. The idealized test was a phishing email sent to all employees and the data acquired provides four fields:

- **UserID**: The same ID's previously constructed on Training and Awareness datasets;
- **Date**: Data regarding the date of test, randomly created between 01.05.2018 and 31.05.2018, the dates are adjusted for each year and semester. The month is selected to be on the subsequent quarter of the training date. The formula used is similar to the one used on the previous dataset:
  `=+RANDBETWEEN(DATE(2018;5;1);DATE(2018;5;31))`
- **Followed**: A random population of 18% of user ID's were selected as following the phishing email and filled with a value of `1`;
- **Reported**: this field indicates if a user reported the email to the Information security team. It was randomly selected over the population as the above field in a total of 78% reporting the email. A total of 4 files corresponding to 2 semesters and 2 years were created where the different random populations were created for Reported and Followed fields.

4.6.4. Password Dataset

Albeit belonging to the first set of metrics, the Training and Awareness, a separate set was created to handle the password data. The data consist of two similar datasets corresponding to two types of passwords, i.e., one for the operational system login and another one for a single-sign-on type of login. These datasets consist on the following fields:

- **Index**: which is just a sequential identifier of the data line;
- **Pass_Category**: One table with the value "`SystemSO`", corresponding to the operational system and the other with "`SSO"`, for single sign on system. The aim was to have distinct type to diversify the analysis.
- **Year**: corresponds to the year of the assessment extracted from the following field;
- **Date_assessment**: This field was generated randomly between 06.01.2018 and 30.09.2019 on a weekly basis by adding 7 days to the original date.
- **Avg Length**: the average number of characters of the password in each assessment, randomly generated between 8 and 32 characters:
  ```
  =+RANDBETWEEN(8,32)
  ```
- **Complexity**: This fields identify the complexity of the assessed passwords in a range between 0 and 100. It was calculated using a normal distribution with a mean of 75 and a standard deviation of 10 by using Excel's formula:
  ```
   =+NORMINV(RAND(),75,5)
  ```
- **Total_Pass**: The total number of passwords assessed in the corresponding date, randomly calculated between 2006 and 2600 (2006 is the minimum number of needed passwords to guarantee logins to the 2006 users):
  ```
  =+RANDBETWEEN(2005,2600)
  ```
- **Crackable**: The number of passwords that were cracked in each assessment, a sample of 10% were calculated using
  ```
  =+INT((RANDBETWEEN(0,10)/100)*H2)
  ```
- **Total_Pass_Not_Compliance**: The total number of passwords that do not comply with the organization's rules. A sample of 8% was selected using a similar expression to the previous one, plus the total of the previous field was added to the final calculation (a *crackable* password is also not compliant):

4.6.5. Vulnerability Dataset

The aim of this dataset is to represent the acquired data from several incidents on information systems. Seven datasets were built to simulate the quarters across 2018 to 2019, where 2019 is composed of only 3 semesters, until September.

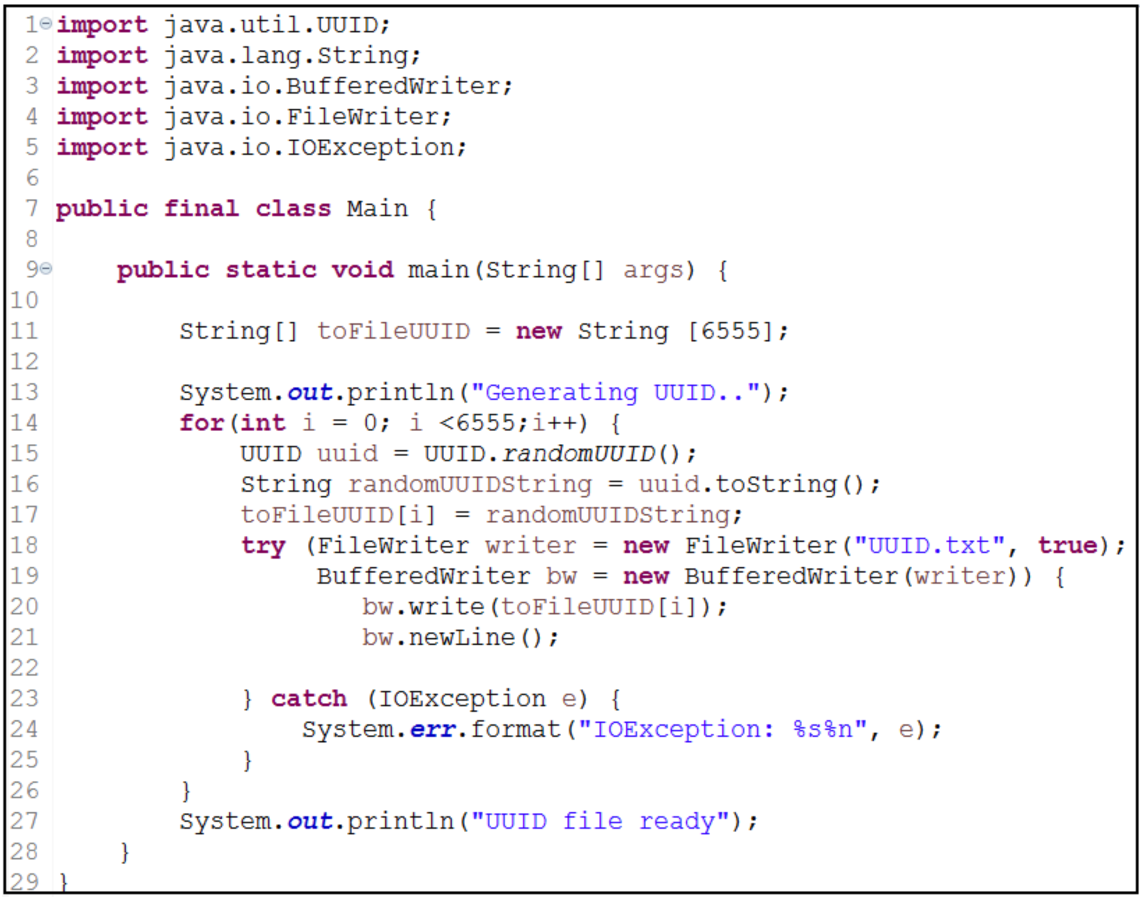
A total of six fields were constructed to populate the tables:

- **UUID_Asset_ID**: This field provides the unique identifier of the information asset. In the absence of a fast method to create the id's, a small Java program was written to randomly generate a total of 6555 distinct UUID's as presented in figure 15.
- **BIA Level**: Represents the Business Impact Analysis level of each assets. For level 3, a sample of 5% were randomly selected, level 2 was a 5% sample from the remaining population and BIA level 1 is the remaining population without the 2 first sample:

  ```
  =+IF(ROW()>5%*COUNTA($A$2:$A$6556);"";INDEX($A$2:$A$6556;RANK.AVG(B2;$B$2:$B$6556;0);1))
  ```
- **Assessment**: This field has a value of 0 or 1, in which the value 1 represents "Assessment made". A sample of 78% was randomly selected:

  ```
  =+IF(ROW()>78%*COUNTA($A$2:$A$6554);"";INDEX($A$2:$A$6554;RANK.AVG(B2;$B$2:$B$6554;0);1))
  ```
- Date Assessment: This is the date on which the assessment was made. The first set was randomly populated between 01.02.2018 and 27.02.2018.
- Patch: When an asset is patched, this field is populated with a value of 1, 0 if no patch was applied. The population was created on a random basis of 83% of the assets that were subjected to assessment.
- Date Patch: A random generated date between 01.03.2018 and 31.03.2018 on the Patch population;

All the seven files follow the same creation rules, except on the percentage of the random populations created as a mean to provide variation in the analysis on the Dashboard.



```
import java.util.UUID;
import java.lang.String;
import java.io.BufferedWriter;
import java.io.FileWriter;
import java.io.IOException;

public final class Main {

    public static void main(String[] args) {

        String[] toFileUUID = new String [6555];

        System.out.println("Generating UUID..");
        for(int i = 0; i <6555;i++) {
            UUID uuid = UUID.randomUUID();
            String randomUUIDString = uuid.toString();
            toFileUUID[i] = randomUUIDString;
            try (FileWriter writer = new FileWriter("UUID.txt", true);
                 BufferedWriter bw = new BufferedWriter(writer)) {
                    bw.write(toFileUUID[i]);
                    bw.newLine();

            } catch (IOException e) {
                System.err.format("IOException: %s%n", e);
            }
        }
        System.out.println("UUID file ready");
    }
}
```

Figure 15 - Java program written to produce the UUID

4.6.6. Incident Handling Dataset

For the purpose of creating information to provide analysis over the incident topic, a dataset from TU[8] Research Data from TU Delft Library[9], was used. This dataset is open to public usage, further details about the general terms of use are presented in Appendix B.

This dataset contains several fields, but for the purpose of this work only the following fields were extracted on the application:

- **Service Component WBS (aff);**
- **Incident ID;**
- **Status;**
- **Impact;**
- **Urgency;**
- **Priority;**
- **Category;**

[8] Downloaded from https://data.4tu.nl/

[9] Technische Universiteit Delft - https://www.tudelft.nl/

- **KM number;**
- **Alert Status;**
- **# Reassignments;**
- **Open Time;**
- **Reopen Time;**
- **Resolved Time;**
- **Close Time;**
- **Handle Time (Hours)**
- **Improvements**

All records where of type field "`incident`". All other non "`incident`" type were removed. All dates fields were modified to provide dates according to the previous data, as such, the year `2012` was changed to `2018`, `2013` to `2018` and `2014` to `2019`, the field Improvements were randomly generated to supply data to "Actions for security improvements" metric.

4.6.7. Assets Inventory Dataset

The vulnerability dataset is strongly related with the inventory dataset, as such and for simplicity, the vulnerability dataset was extended with two more fields for the purpose of this metrics group, therefore on the last vulnerability dataset two more fields were added and this group was considered as an annual assessment:

- **IP Address,** The IP address[10] of a given asset. This field was randomly generated using web application at `https://onlinerandomtools.com/generate-random-ip` in a total of 7131 distinct addresses vs 6553 assets to enable the simulation of unauthorized devices (those IP Addresses that do no map to an asset UUID).
- System Name: A name that is an attribute of an asset. Randomly calculated using a hard-coded component defined as 'SYS0' plus a randomly calculated 6 characters part over an abecedary:

```
=+CHAR(RANDBETWEEN(65;90))&CHAR(RANDBETWEEN(65;90))&CHAR(RANDBETWEEN(65;90))&CHAR(RANDBETWEEN(65;90))&CHAR(RANDBETWEEN(65;90))&CHAR(RANDBETWEEN(65;90))
```

[10] IP address is an Internet Protocol address that identifies any device connected to a computer network using Internet Protocol.

Chapter 5
Design of Dashboard Application

This chapter aims at explaining and presenting the design and development process focusing on exposing the black box side of the application.

5.1. Construction of Information Security Dashboard

This section explains the construction of the final product, the Cyber Situational Awareness Dashboard for Information Security. It outlines and explains the construction of the data model and the building of the front end. Figure 16 provides a general overview of the development of the CSADIS application where the functional structure approach is displayed.

5.1.1. Data model

The construction of the data model represents the first step of the development of the application. A business intelligence application shall be sustained by a dimensional data model. It is simpler and cleaner than a usual relational data model, it is denormalized, it represents a business process analysis and it is aimed on data analysis and reporting, which is the ultimate goal of a BI tool. For a Qlikview/Qlik Sense implementation, a star schema data model is the recommended option as it is optimizes performance on the in-memory data engine, the Qlik QIX engine.

By using such model, it becomes easier to understand the conceptualized business model and, above all, because it performs better and more efficiently as it avoids unnecessary joins between several tables (Kimball & Ross, 2013).

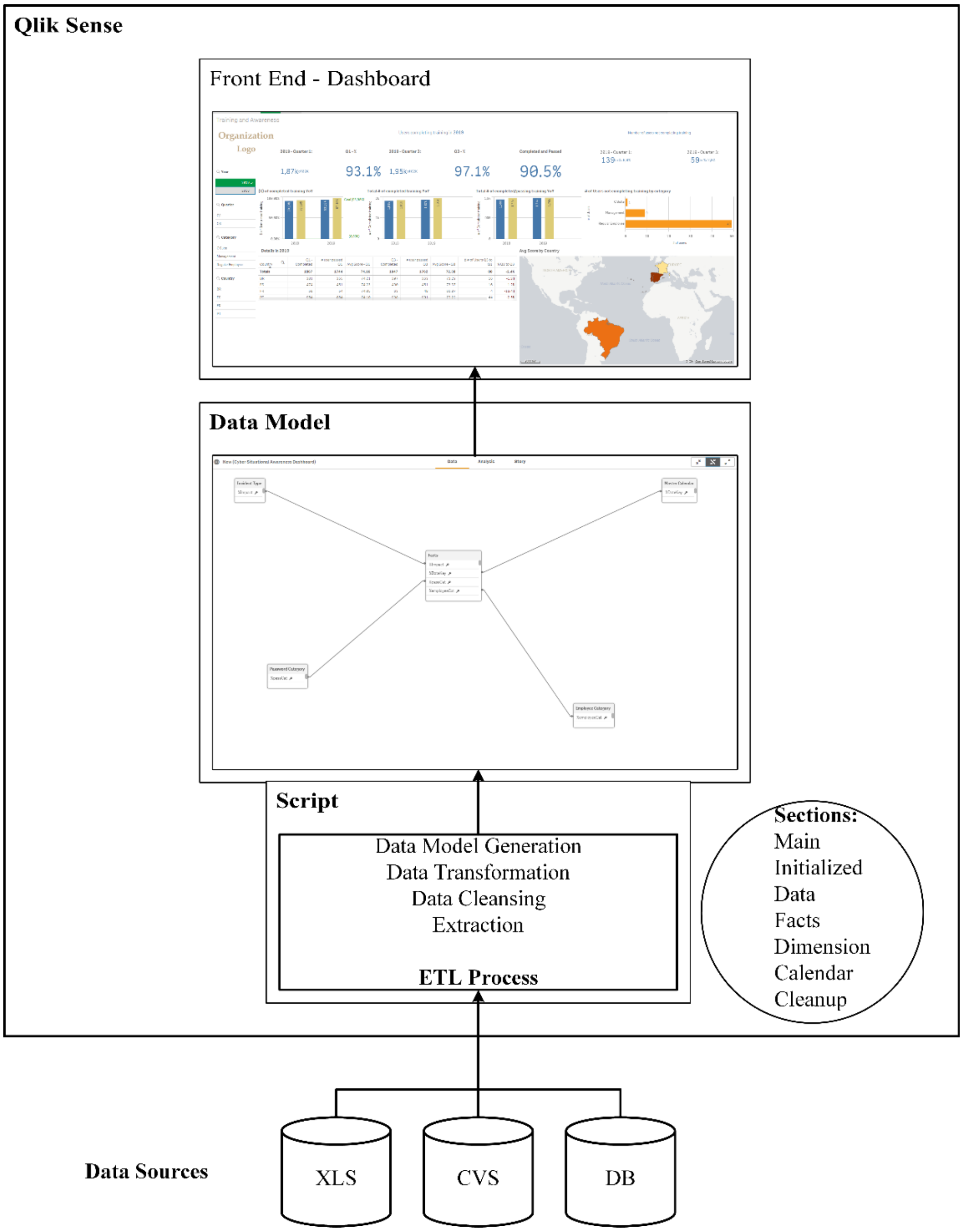
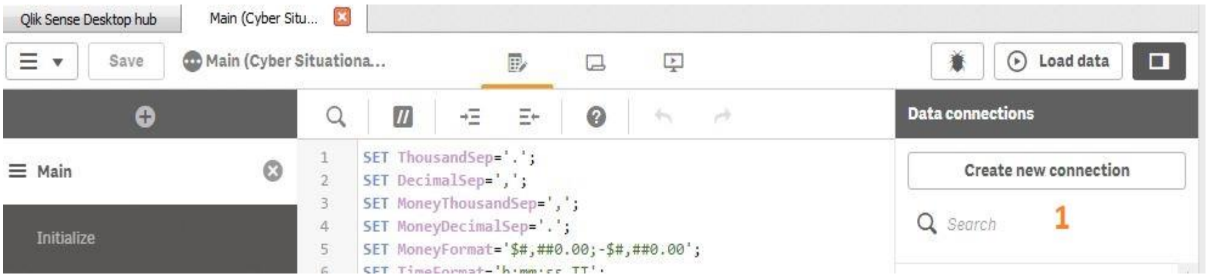


Figure 16 - Dashboard development approach

Loading tables

Qlik Sense allows manual or automatic generation of the script. For the purpose of this application, manual generation was chosen as it allows a more controlled and customized way of work for complex scripts. The first step to build the data model is to load the tables into the application. To accomplish this, a connection string to the data was created, as shown on figure 17.



```
SET ThousandSep='.';
SET DecimalSep=',';
SET MoneyThousandSep=',';
SET MoneyDecimalSep='.';
SET MoneyFormat='$#,##0.00;-$#,##0.00';
SET TimeFormat='h:mm:ss TT';
```

Figure 17 - Qlik Sense Connection creation

The present work is only constructed with excel files that simulate tables or views in a database, therefore, a connection is created to a file location, to a specific folder where the XLS files were previously placed. Although the current application only works with files, Qlikview and Qlik sense allow to connect different types of data sources, such as: cloud services, regular data bases (ODBC, OLE DB), SAP systems, web services, and others. The dialogue to add connections to the application is shown on Figure 18.

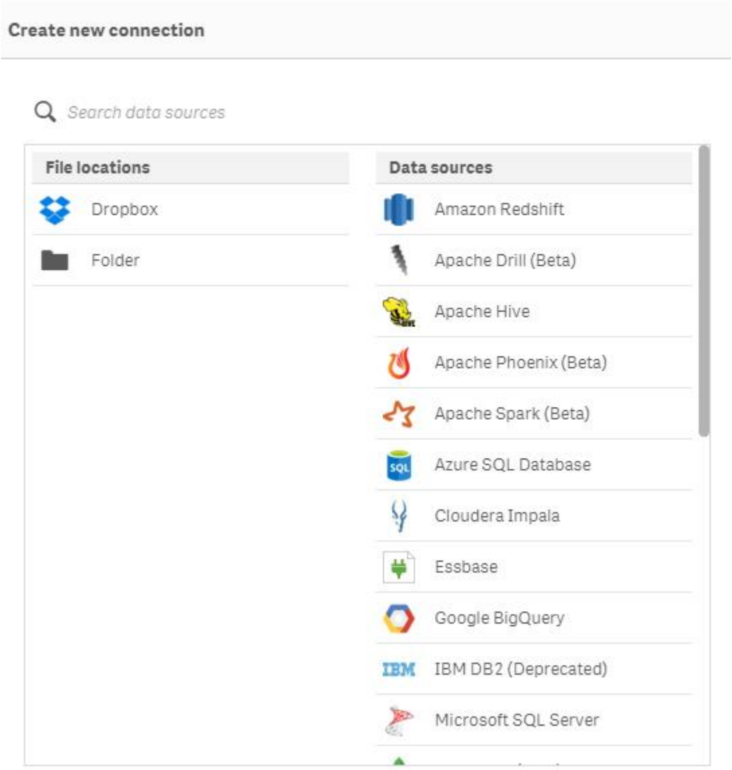
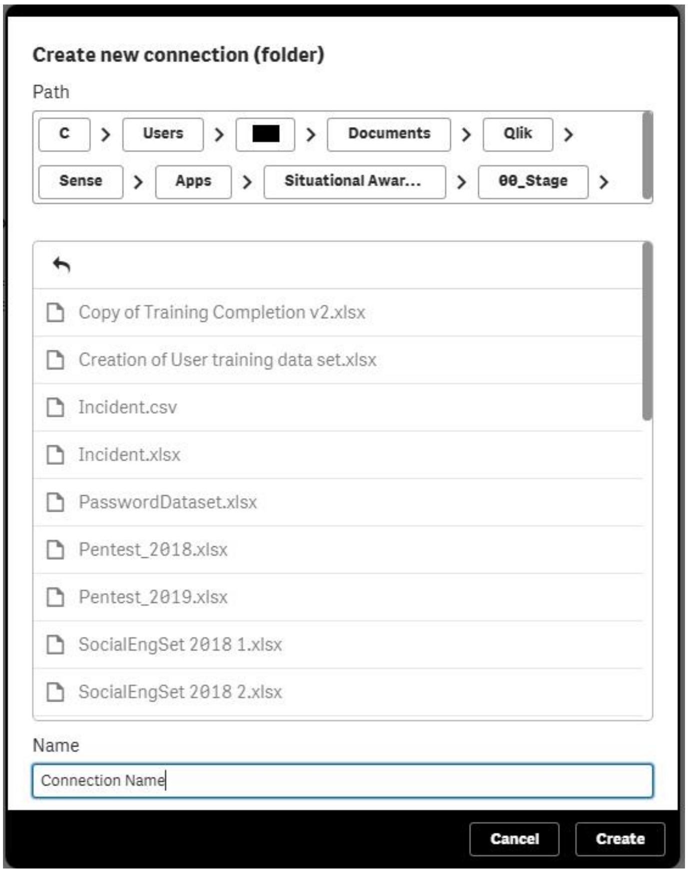


Figure 18 – Creating data connections

Figure 19 shows the creation of a connection to a file system, which is what was done in this work. After a connection to a folder is created, the developer is able to access all the files it contains.



Create new connection (folder)

Path

C > Users > [redacted] > Documents > Qlik > Sense > Apps > Situational Awar... > 00_Stage >

- Copy of Training Completion v2.xlsx
- Creation of User training data set.xlsx
- Incident.csv
- Incident.xlsx
- PasswordDataset.xlsx
- Pentest_2018.xlsx
- Pentest_2019.xlsx
- SocialEngSet 2018 1.xlsx
- SocialEngSet 2018 2.xlsx

Name

Connection Name

Figure 19 - Connection to a file system

Data Modelling

First, seven distinct script sections were created on the script to organize the various components of code, as can be seen on the left side of Figure 17. Sections are only code separators which help organize the scripting section and no naming convention exists. The names hereby introduced are merely to help navigate and split the code.

On the current application, seven coding sections were created:

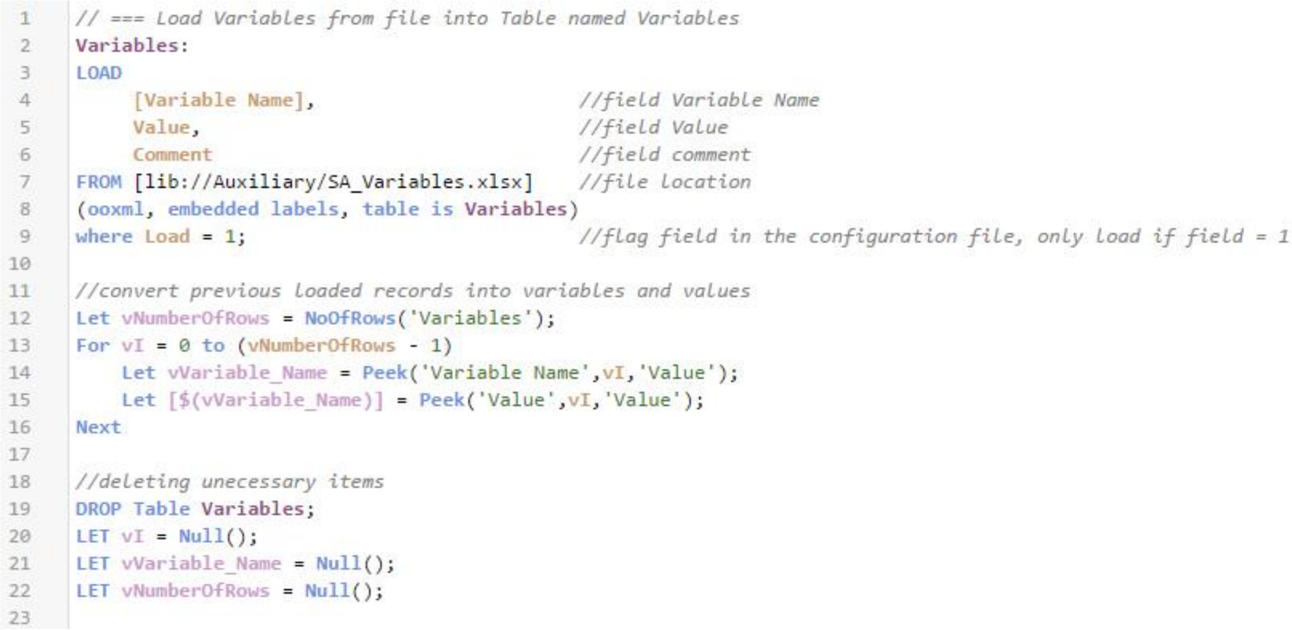
- Main;
- Initialized;
- Data;
- Facts;
- Dimension;
- Calendar;
- Cleanup.

The *Main* section contains the Qlik definitions automatically created for date and time formats, decimal separators, and other regional settings. The *Initialize* section is where an external file was loaded with the variables that contain the set analysis expressions that will be used in the front-end objects. Although not mandatory, loading expressions from an external file allows to make any maintenance or changes in a centralized manner, therefore in a single point. Whenever there is the need to add any changes, only one place has to be changed and, as long as the objects have the correct variable name in the expression's editor, these objects will respond to any changes that are introduce. Hence, any change is less error prone and provides faster development. An XLS file named SA_Variables.xlsx was created with four fields:

- Variable Name: the name of the variable to be used in the application;
- Value: the content is an expression to be assign to the variable;
- Comment: in case any developer comments shall be added;
- Load: a flag field where the value 1 denotes that the corresponding record is to be loaded or, value 0, not loaded.

As mentioned in Section 3.2.1, Qlik uses a native scripting language. This language allows for data manipulation, similar to SQL, but extends to other programming languages common operations and functionalities, such as flow control statements and subroutines.

Figure 20 shows the loading of the variable configuration file and subsequent transformation of the data into variables in the applications.



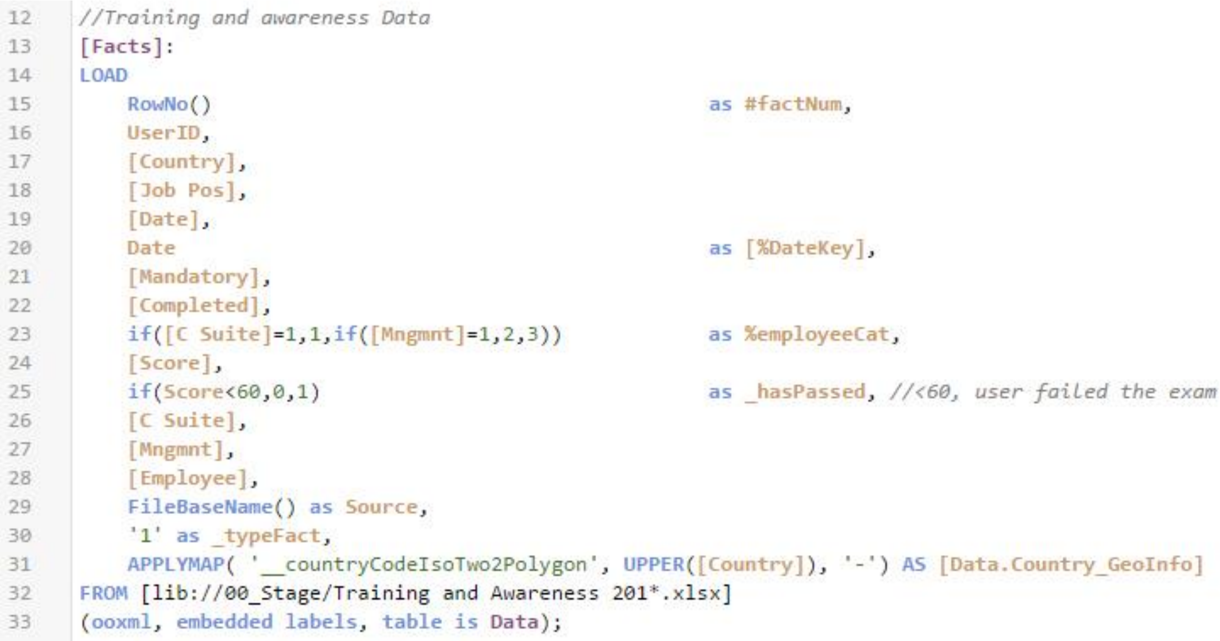
```
// === Load Variables from file into Table named Variables
Variables:
LOAD
     [Variable Name],                       //field Variable Name
     Value,                                 //field Value
     Comment                                //field comment
FROM [lib://Auxiliary/SA_Variables.xlsx]    //file location
(ooxml, embedded labels, table is Variables)
where Load = 1;                             //flag field in the configuration file, only load if field = 1

//convert previous loaded records into variables and values
Let vNumberOfRows = NoOfRows('Variables');
For vI = 0 to (vNumberOfRows - 1)
    Let vVariable_Name = Peek('Variable Name',vI,'Value');
    Let [$(vVariable_Name)] = Peek('Value',vI,'Value');
Next

//deleting unecessary items
DROP Table Variables;
LET vI = Null();
LET vVariable_Name = Null();
LET vNumberOfRows = Null();

```

Figure 20 - External configuration file with variables to be loaded into the application

The *Data* section, which is related with geographic information, was automatically generated by the application to load the necessary data related with the use of wold maps visualizations. The code was automatically created in accordance with the data manually loaded in the fact tables. The dataset 'training and awareness' differentiates the users by countries, therefore Qlik Sense generated the necessary data from its maps extension to connect to the ISO country codes[11] provided in the dataset of the application. This section will not be shown in this work as it was automatically generated by Qlik Sense.

The section *Facts* is where all the previously created datasets were loaded into the application and where was performed the necessary transformations steps to the data. These ETL processes were applied both to the Facts and Dimensions sections. Figure 21 show a table being loaded into the application. Line 13 is where the name of the table is created, for subsequent tables loaded into the application's fact table, a concatenate operation would be done, similar to SQL's append, as shown on line 89 on Figure 22, the rest of the code is simply regular data loading operations.

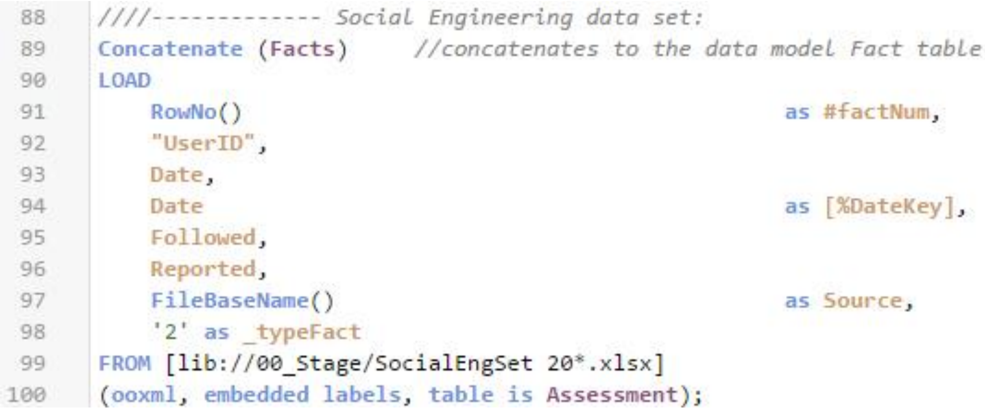
An example of data transformation is on lines 117, 118 and from 121 to 123 of Figure 23, where transformation steps of the date type at the Incident dataset can be seen. The date and time fields, were on an unusual format, requiring some formatting process (data transformation process) to be in accordance to the previously date formats in use.



```
12  //Training and awareness Data
13  [Facts]:
14  LOAD
15      RowNo()                                         as #factNum,
16      UserID,
17      [Country],
18      [Job Pos],
19      [Date],
20      Date                                            as [%DateKey],
21      [Mandatory],
22      [Completed],
23      if([C Suite]=1,1,if([Mngmnt]=1,2,3))            as %employeeCat,
24      [Score],
25      if(Score<60,0,1)                                as _hasPassed, //<60, user failed the exam
26      [C Suite],
27      [Mngmnt],
28      [Employee],
29      FileBaseName() as Source,
30      '1' as _typeFact,
31      APPLYMAP( '__countryCodeIsoTwo2Polygon', UPPER([Country]), '-') AS [Data.Country_GeoInfo]
32  FROM [lib://00_Stage/Training and Awareness 201*.xlsx]
33  (ooxml, embedded labels, table is Data);
```

Figure 21 - Creating the Facts table

[11] ISO 3166 country codes is a standard of two letters that represent each country name.

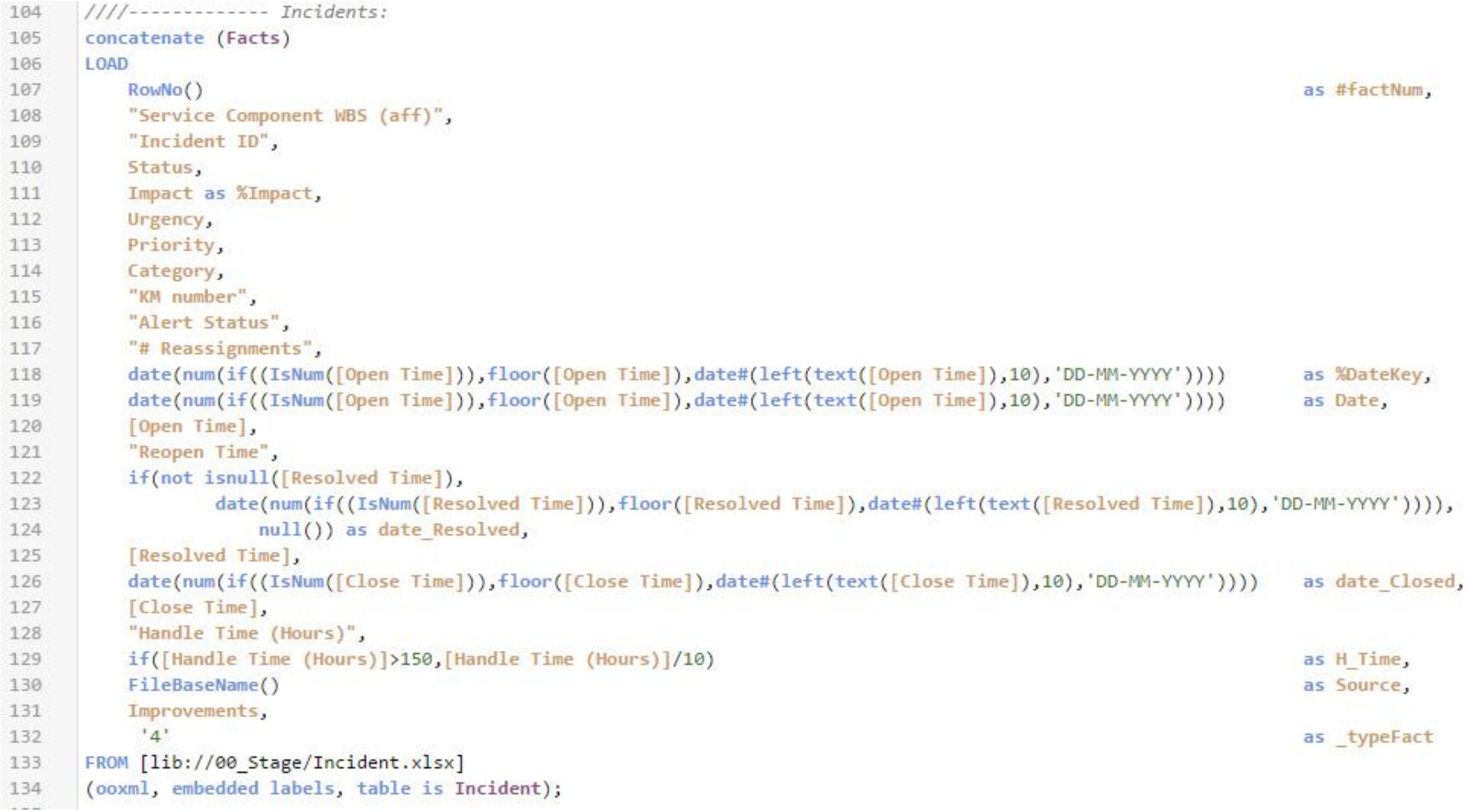


```
////------------ Social Engineering data set:
Concatenate (Facts)     //concatenates to the data model Fact table
LOAD
    RowNo()                                         as #factNum,
    "UserID",
    Date,
    Date                                            as [%DateKey],
    Followed,
    Reported,
    FileBaseName()                                  as Source,
    '2' as _typeFact
FROM [lib://00_Stage/SocialEngSet 20*.xlsx]
(ooxml, embedded labels, table is Assessment);
```

Figure 22 - Concatenation of a table to the application's Fact table

All loaded tables into the central Fact table share 4 common control fields defined in the script:

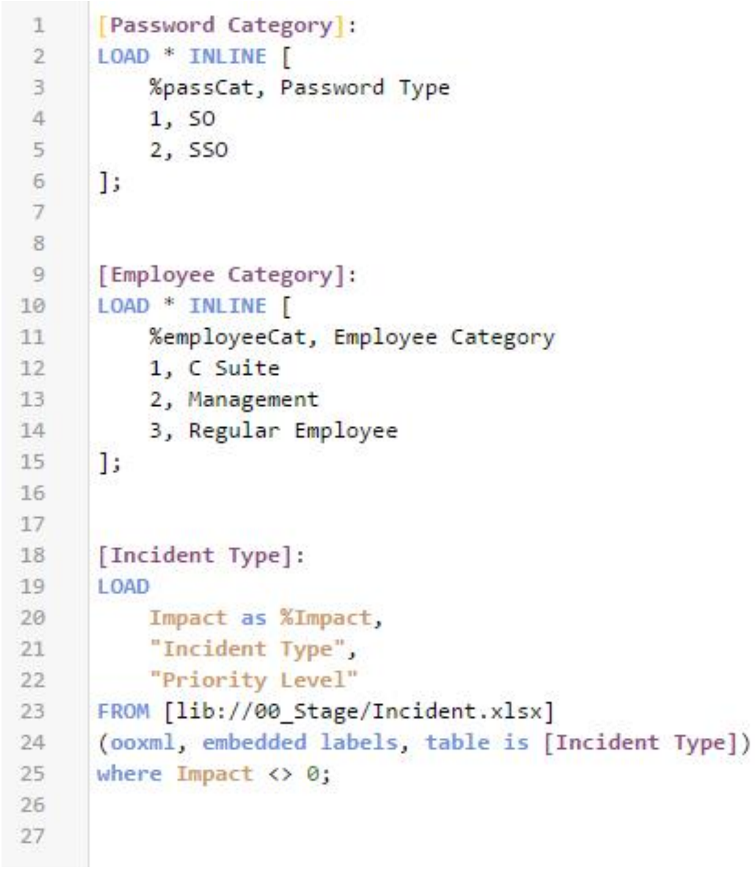
- `#factNum`: a sequential identifier for each record loaded;
- `%DateKey`: a key field to connect the fact to a calendar;
- `Source`: identifier of the source file loaded into the table;
- `_typeFact`: Fact Id which helps identify in the set analysis expressions in the frontend, with the following values:
  - 1 - Training and Awareness
  - 2 - Social Engineering
  - 3 - Password Complexity
  - 4 - Incident
  - 5 - Vulnerability
  - 6 - Pentest
  - 7 - Assets Inventory



```
////------------- Incidents:
concatenate (Facts)
LOAD
    RowNo()                                                                                                 as #factNum,
    "Service Component WBS (aff)",
    "Incident ID",
    Status,
    Impact as %Impact,
    Urgency,
    Priority,
    Category,
    "KM number",
    "Alert Status",
    "# Reassignments",
    date(num(if((IsNum([Open Time])),floor([Open Time]),date#(left(text([Open Time]),10),'DD-MM-YYYY'))))       as %DateKey,
    date(num(if((IsNum([Open Time])),floor([Open Time]),date#(left(text([Open Time]),10),'DD-MM-YYYY'))))       as Date,
    [Open Time],
    "Reopen Time",
    if(not isnull([Resolved Time]),
            date(num(if((IsNum([Resolved Time])),floor([Resolved Time]),date#(left(text([Resolved Time]),10),'DD-MM-YYYY')))),
                null()) as date_Resolved,
    [Resolved Time],
    date(num(if((IsNum([Close Time])),floor([Close Time]),date#(left(text([Close Time]),10),'DD-MM-YYYY'))))    as date_Closed,
    [Close Time],
    "Handle Time (Hours)",
    if([Handle Time (Hours)]>150,[Handle Time (Hours)]/10)                                                   as H_Time,
    FileBaseName()                                                                                          as Source,
    Improvements,
     '4'                                                                                                    as _typeFact
FROM [lib://00_Stage/Incident.xlsx]
(ooxml, embedded labels, table is Incident);
```

Figure 23 - Concatenation of Incidents data

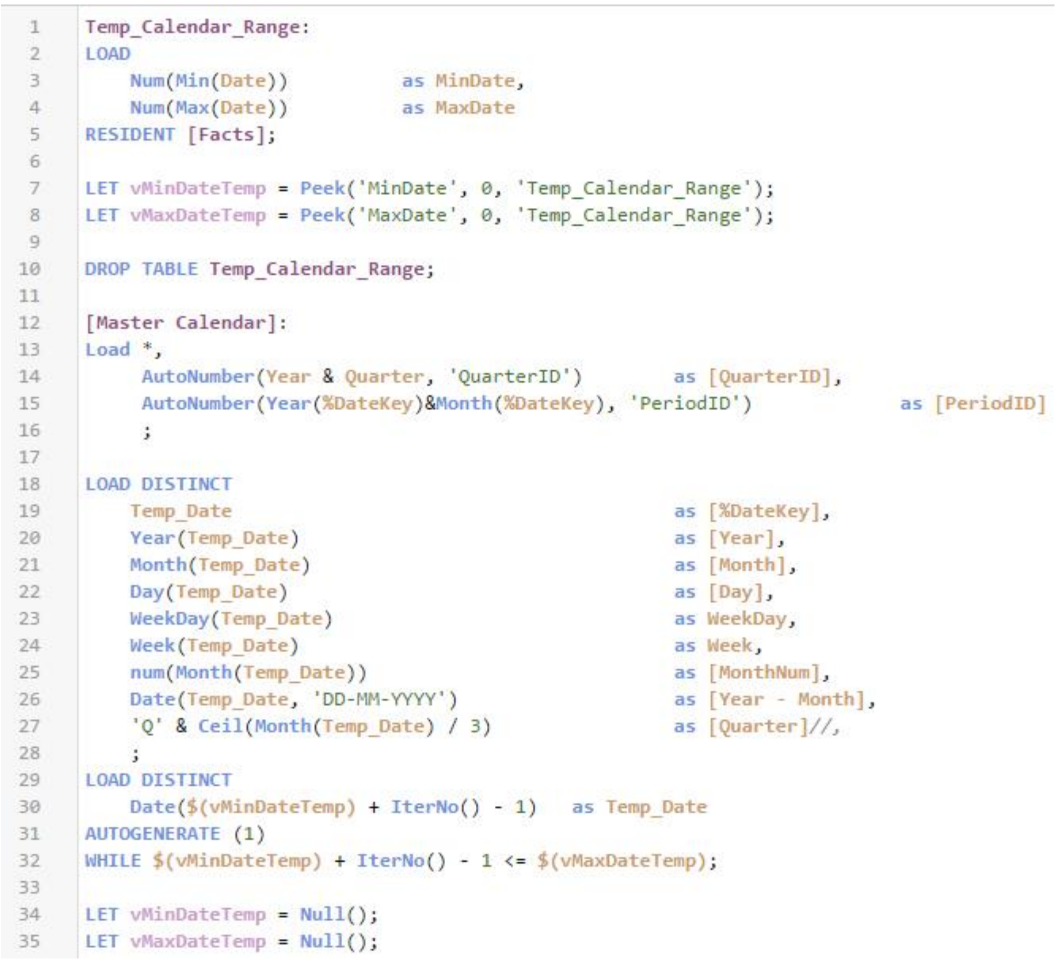
The next section is the *Dimension* section and it is where the dimension tables are loaded. As can be seen on Figure 24, in this work, three tables were created in which the categorization of employee, password and incident are defined and connected to the fact table. These dimensions contain the description of each category (alphanumeric data type) and thus, avoiding repetitions in the fact table. They are connected by a key field to the *Facts* table with a numeric data type, besides the theory in dimensional modelling defining that attributes of facts shall be on dimension tables, this is also important because an integer type occupies less memory than an alphanumeric type in the Qlik engine.



```
[Password Category]:
LOAD * INLINE [
    %passCat, Password Type
    1, SO
    2, SSO
];


[Employee Category]:
LOAD * INLINE [
    %employeeCat, Employee Category
    1, C Suite
    2, Management
    3, Regular Employee
];


[Incident Type]:
LOAD
    Impact as %Impact,
    "Incident Type",
    "Priority Level"
FROM [lib://00_Stage/Incident.xlsx]
(ooxml, embedded labels, table is [Incident Type])
where Impact <> 0;
```

Figure 24 - Dimensions

On the Calendar section, a master calendar was developed to allow temporal analysis of the data. It contemplates all the dates of the F*acts*, from the minimum to the maximum date and it is connected to the Facts table by a key field denominated `%DateKey`, as demonstrated on Figure 25.



```
Temp_Calendar_Range:
LOAD
    Num(Min(Date))          as MinDate,
    Num(Max(Date))          as MaxDate
RESIDENT [Facts];

LET vMinDateTemp = Peek('MinDate', 0, 'Temp_Calendar_Range');
LET vMaxDateTemp = Peek('MaxDate', 0, 'Temp_Calendar_Range');

DROP TABLE Temp_Calendar_Range;

[Master Calendar]:
Load *,
     AutoNumber(Year & Quarter, 'QuarterID')        as [QuarterID],
     AutoNumber(Year(%DateKey)&Month(%DateKey), 'PeriodID')            as [PeriodID]
     ;

LOAD DISTINCT
    Temp_Date                                       as [%DateKey],
    Year(Temp_Date)                                 as [Year],
    Month(Temp_Date)                                as [Month],
    Day(Temp_Date)                                  as [Day],
    WeekDay(Temp_Date)                              as WeekDay,
    Week(Temp_Date)                                 as Week,
    num(Month(Temp_Date))                           as [MonthNum],
    Date(Temp_Date, 'DD-MM-YYYY')                   as [Year - Month],
    'Q' & Ceil(Month(Temp_Date) / 3)                as [Quarter]//,
    ;
LOAD DISTINCT
    Date($(vMinDateTemp) + IterNo() - 1)   as Temp_Date
AUTOGENERATE (1)
WHILE $(vMinDateTemp) + IterNo() - 1 <= $(vMaxDateTemp);

LET vMinDateTemp = Null();
LET vMaxDateTemp = Null();
```

Figure 25 - Calendar generation

The last section, *Cleanup*, is composed of maintenance activities, as for example, deleting unneeded variables or unused fields.

The creation of key fields in Qlikview/Qlik Sense is a straightforward process, if two fields in two distinct tables have the same name, Qlik will create a relationship between the two fields. Hence, these fields will be the key fields that connect two or more tables. In Qlik, it is expected to have two tables connected by only one field. Therefore, if two tables share more than one field, this will create a synthetic key, a problematic situation that will increase memory consumption.

If more than two tables are connected in such a way that several paths of associations between the data is generated, a circular reference is created. This situation leads to possible memory increase and ambiguity in the data relationship. Synthetic keys and circular references shall be avoided at all cost, by using a composite key as a workaround solution in the first situation and by renaming field's names for the second one. By implementing only one common field between each table, a bidirectional connection is obtained through every table in the model, this is essential to the Qlik's associative model to work correctly. By doing so, Qlik implementation allow a dynamically analysis of all the data, without having to aggregate the data priory, which would limit the data analysis on the front end to previously idealized models. Following the best practices rules for Qlik development, all key fields starts with the percentage "%" character and all flag fields start with underscore '_', this is illustrated in figure 26.

As previously stated, this application uses heterogeneous data from distinct sources that mimics different data bases, plus, the data has different granularity. To implement the data model in this application, one fact table was produced in order to obtain a star schema dimensional model. Hence, the fact table of this application is the result of concatenating multiple distinct facts, that is, the distinct datasets for each topic, into one unique table.

Another solution would be to use multiple fact tables connected through a link table, however the dimensional model obtained through that solution would be closer to a snow flake schema. While it would not be an incorrect implementation, it wouldn't be the optimal solution. To overcome this hindrance, a flag field called `_typeFact` was used which contains a number code previously defined for each fact loaded, as explained earlier. Using this approach, every set analysis expression starts by setting the flag field to the desired fact that is being analysed at the front end. With careful planning, detailed set analysis coding and reviewing of the differences in time granularity of each fact, this solution proved to be sufficient in this situation, as it provides complex and associative data analysis as desired. However, for far more complex implementations, a step forward shall be considered as for example, the use of generic keys and/or the implementation of multiple calendars that would allow for more complex temporal analysis and time correlations for greater amounts of facts. The specificities of these advanced models are beyond the scope of this work, in any case, if necessary, there are several solutions that can be implemented using Qlikview/Qlik Sense. Figure 26 demonstrates part of the final data model implemented in the application using a star schema model as previously proposed.

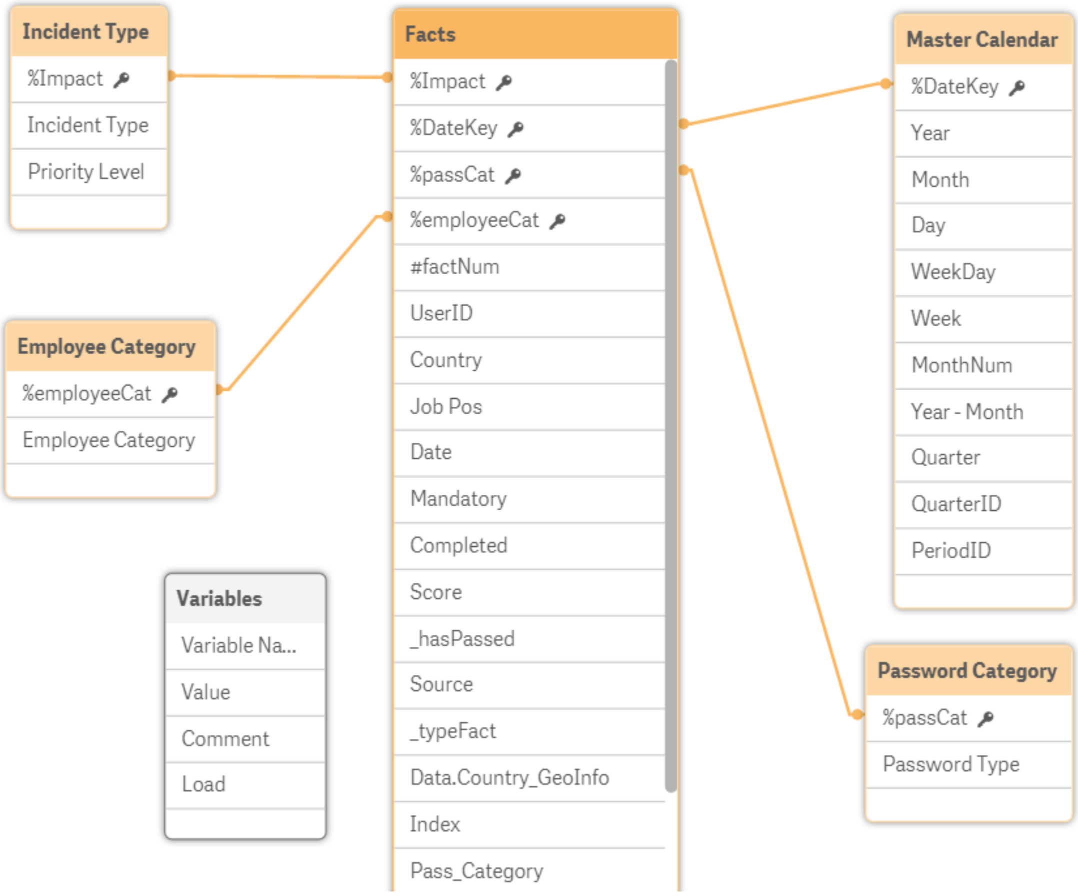


Figure 26 - Application's star schema data model

5.1.2. Development of The Front End

At the front-end part is were all metrics and objects were constructed to provide information analysis and answers about the security state.

The first page of the application is related with the Training and Awareness metrics, the second with the Social Engineering topics, Password strength is the third page, the Vulnerability is presented at the fourth, at the fifth page is Incident Handling and, finally, the assets inventory analysis is provided at the sixth page of the application.

At this level is where the Qlik associative model is perceived. Whenever the user selects any information, or *clicks* in anything shown on the front end, the engine is creating a subset of data that filters out all the information in the front end according to a user defined selection state of the data, and all the visual objects would react to this as a filter, except if encoded at a set analysis level to not respond to all or certain dimensions. This allows for a dynamic analysis and exploration of data where the limit is only the data previously loaded in the script.

The first step is to create a new sheet where all the analysis objects will be place. Then, each object, such as charts, text boxes, shall be placed accordingly to the detail level, from left to right, top to bottom. Image 27 portraits the development module of Training and Awareness Dashboard. On the left side, Detail Z, is where the type of objects are selected and dragged to the visualization sheet, shown on detail X. On Figure 27, detail A, the KPI *Users completing training in* `[year]` is shown, and on detail B, is the object properties where the set analysis expression, that is, the metric, is developed, as well as all other design objects properties, such as colour, font dimension, number format and so on. For all the objects in the dashboards on this application, this workflow applies. Table 4 provides all the set analysis expressions used in this dashboard. On figure 27, object A used the set analysis expressions depicted in Table 4 as 5.1 and 5.2 to provide the total number of users completing the training sessions in the selected year, for the first quarter, and the total number of users at the analysed date. Objects in detail C and E, show the number of users completing the training, but on a percentage basis. The corresponding metrics are shown on Table 4, expression number 5.5 and 5.6. Object presented in detail D is similar to the object in detail A, but for the latest quarter and the expressions are presented with the corresponding numbers 5.3 and 5.4.

The KPI object F shows the percentage of users that completed the training and passed the exam, and the metric expressions is shown as set analysis expression in Table 4 with the number 5.15. KPI objects presented in details G and H provides the absolute and relative values of users not completing the training on first and second this quarter of the selected year and their expressions are 5.10, 5.11 for detail G and 5.12 and 5.13 for detail H.

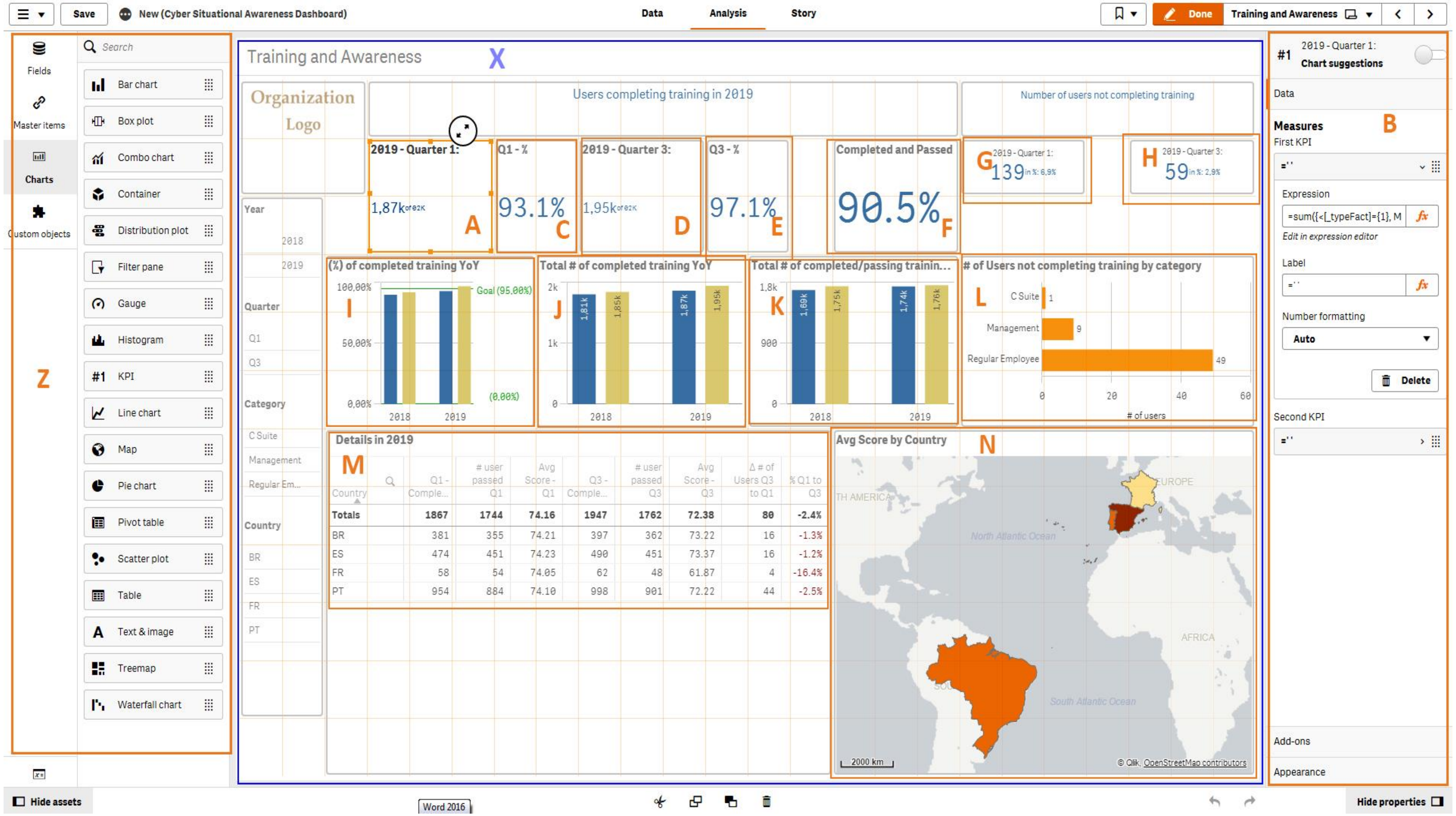
Bar chart objects portrayed at detail I, J and K provide a comparison between the current and prior year. The metrics for these objects is presented in Table 4, expressions 5.7, 5.8 and 5.9. Object I also provide a threshold goal line, to provide the user with a fast overview of completeness of the metrics. Object depicted in detail M is of type table, usually used to provide further details, and provides several metrics for comparison purposed. The metrics presented in this object are defined in table 4 from 5.16 to 5.23. Lastly, the object map, which uses a colour gradient to provide analysis over the average scoring of the exam by country is shown in detail N and the metric used is the number 5.24.

All the expressions created in the context of the development of the Dashboard are presented in Appendix E.

Table 4 - Set analysis expressions developed on Training and Awareness Dashboard

Metric Name	Set Analysis	Number
Total # users Completing Training Q1	sum({<[_typeFact]={1}, Month=, Day=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"} >}Completed)	5.1
Total # users Q1	=' of '&chr(23)&num(count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)/1000,'##.00')&'K'	5.2
Total # users Completing Training Q3	sum({<[_typeFact]={1}, Month=, Day=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"} >}Completed)	5.3
Total # users Q3	=' of '&chr(23)&num(count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)/1000,'##.00')&'K'	5.4
% Users completing training Q1	count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID))	5.5
% Users completing training Q3	count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID))	5.6
% Users completing training YoY	num((sum({<[_typeFact]={1},Year=, Month=, Day=>}Completed)/count({<[_typeFact]={1},Year=, Month=, Day=, Completed={0,1} >}UserID)),'##,#0%')	5.7
# Users completing training YoY	sum({<Year=, [_typeFact]={1}, Month=, Day=>}Completed)	5.8
# Type of users not completing training	count({<[_typeFact]={1}, B25, Year=, Quarter=, Month=, Completed={0} >}distinct UserID)	5.9
# users not completing training Q1	=' in %: '&num(1-(count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)),'##,#%')	5.10
% users not completing training Q1	=' in %: '&num(1-(count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)),'##,#%')	5.11

# users not completing training Q3	=' in %: '&num(1-(count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)),'##,#%')	5.12
% users not completing training Q3	=' in %: '&num(1-(count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)),'##,#%')	5.13
# users not completing training by Category	count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0} >}UserID)	5.14
Users completing and passing training	(sum({<[_typeFact]={1}, PeriodID = {$(vMaxTAPeriod)},Year = {"$(=Year(Max(Date)))"}, Completed={1}>}_hasPassed) /sum({<[_typeFact]={1}, PeriodID = {$(vMaxTAPeriod)},Year = {"$(=Year(Max(Date)))"} >}Completed))	5.15
Table: # Completed Q1	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)	5.16
Table: # Completed Q3	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)	5.17
Table: # Passed Q1	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1}, [_hasPassed]={1} >}UserID)	5.18
Table: # Passed Q3	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1}, [_hasPassed]={1} >}UserID)	5.19
Table: AVG Score Q1	Avg({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}Score)	5.20
Table: AVG Score Q3	Avg({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}Score)	5.21
Table: Δ # of Users Q3 to Q1	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)-count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)	5.22
Table: Δ % Q1 to Q3	(Avg({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}Score)/Avg({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}Score)-1)	5.23
Map: Average Score by country	Avg({<[_typeFact]={1}, Day=, Month=, Quarter=,Year =, PeriodID = {$(vMaxTAPeriod)}, Completed={1} >}Score)	5.24



Details in 2019

Country	Q1 - Comple...	# user passed Q1	Avg Score - Q1	Q3 - Comple...	# user passed Q3	Avg Score - Q3	Δ # of Users Q3 to Q1	% Q1 to Q3
Totals	**1867**	**1744**	**74.16**	**1947**	**1762**	**72.38**	**80**	**-2.4%**
BR	381	355	74.21	397	362	73.22	16	-1.3%
ES	474	451	74.23	490	451	73.37	16	-1.2%
FR	58	54	74.05	62	48	61.87	4	-16.4%
PT	954	884	74.10	998	901	72.22	44	-2.5%

#1 2019 - Quarter 1:
Chart suggestions

Data

Measures

First KPI
='

Expression
=sum({<[_typeFact]={1}, M

Edit in expression editor

Label
='

Number formatting
Auto

Delete

Second KPI
='

Add-ons

Appearance

Figure 27 - Training and Awareness Dashboard Development

This chapter summarizes the creation of the application that demonstrates the goal of this dissertation. Several constraints and difficulties were observed, both on technical level, such as the creation of a fact table based on multi facts and of multi granularity, but also at the tool selected to develop and implement. It was expected to have the application developed in Qlikview, however, Qlik provided only a student license for Qlik Sense tool. Although the scripting language, the syntax language and set analysis are the same, several other structural elements were quite different, especially on the front end and at the objects level, which culminated in a larger learning curve. In Appendix D, a table containing all the metrics transformed into set analysis to be used in the application is provided.

Chapter 6
Discussion

This chapters discusses the outcome of this study, from the literature review on information security metrics and situational awareness to the use of analytics and business intelligence tools for the construction of a dashboard application. Ultimately, the Situational Awareness Dashboard application is where all the topics of this work sum up into a product, which presents situational awareness that enables decision making and provides the drivers for action planning.

6.1. Security Metrics

Based on the literature review it can be concluded that metrics of IS represent a wide field of study that is relevant for both academic and industry communities, composed of several options and implementations. It proves to be difficult to accomplish in a one-time effort, but rather in a continuous improvement and development model based on periodic reviews of new necessities and on the information technology systems involved. This goal can be achieved by reviewing current metrics and understand the correctness and the usefulness of such implemented measures, by assessing the value derived from such implementations versus the effort to develop and maintain it.

From the literature review, it was observed that several studies and discussions were conducted, however, most of the authors refer to the *de facto* standards and frameworks as the ISO/IEC 27000 series or the NIST 800. Other guides have been progressing widely as the CVSS for vulnerabilities or the CIS controls and its associated set of metrics, both having its development relying on the security professional’s community around the globe. During this study, several metrics were chosen on chapter 4 and defined in section 4.4. Some were directly implemented and others were extended in the final step of this work, the dashboard.

6.2. Implementation of metrics

An outcome from this study is that the implementation of a metrics program is a never-ending process. As such, to accomplish two main objectives, securing information systems and getting return of investment, an incremental implementation shall be followed. By having such an incremental process, an organization can quickly deploy an initial set of metrics, thus obtaining immediate results, and by keeping a continuous improvement process, the organization can move forward into a more complex and mature security model.

This was the premise to implement the dashboard. It would be impossible to implement all the metrics gathered during this study and as the final goal of this study was to develop a dashboard, options had to be taken, as such, several metrics were chosen to be implemented, based on the capacity to recreate customized data or the probability to access open datasets to answer the necessities of each metric.. Besides, the degree of importance assigned to each metric is directly related with each organization needs and goals. Thus, the aim of the development was to document the implementation of a security dashboard, that provides relevant information as an outcome and allows dynamic visual analysis, that value can be obtained with an initial set of metrics, and that by keeping a holistic overview, a continuous improvement can be applied culminating in an increased maturity level.

6.3. Situational Awareness and continuous monitoring

Situational awareness and continuous monitoring are intrinsically connected. A continuous monitoring of the information systems controls with a strong metrics program implemented, provides the means for a situational awareness of an organization. To keep on the edge of the information relevance, a continual monitoring of the capabilities and importance of the controls and metrics shall be performed. Such model can be used with a Plan-Do-Check-Act (PDCA) method, mostly for internal analysis, or based on an OODA loop which provides the methodology for a fast adaptation to external actions in contexts of uncertainty and incompleteness. Whatever the methodology used, or even combining them, the goal is to advance the model into a maturity level that provides information superiority.

The development of the dashboard is based on this principle, by providing not only the necessary information for action, but by being developed with this idea in mind, it is ready to add disparate data as needed to accommodate new information needs.

6.4. Dashboard General Overview

This section provides an overview of the several sheets designed and implemented in the application that compose the CSADIS dashboard. It aims to describe the implemented dashboards and the information provided in each one. For text formatting purposes, all Dashboards images are available at the end of this section,

The final goal of this study was to develop a Situational Awareness dashboard that was based on the establishment and implementation of an initial set of metrics that could provide extensibility for future implementations. Information systems differ widely and provide different types of data, by using a business intelligence and analytic tool like Qlik Sense, the possibility to interact with heterogeneous data types and to connect with multiple types of data sources was accomplished.

6.4.1. Training and Awareness analysis

On this dashboard, displayed as Figure 27 in the previous chapter, the metrics discussed at 4.2.1 were introduced and extended, as well as several objects that provide a visual analysis. For comparison purposes, instead of just showing the current state of a given metric, the information of the previous assessment date is also provided. For example, training completeness, being a semester assessment, is shown for both the end of the first and the third quarter, thus providing an overview of the evolution between dates. The metrics were also built to provide absolute figures and relative numbers, even though the metrics defined previously are usually defined in one way only. This is because, with relative numbers it is simpler to compare with a threshold, however with an absolute value the user gets a direct overview of the total figures. In real life applications it would be discussed with the stakeholders and, probably, the presentation would be done in one way only as this could be considered superfluous and a time-consuming process resulting in a waste of resources.

This dashboard, *Training and Awareness*, demonstrates metrics for training information security statistics such as the number and percentage of users completing the training in the selected year– by default the current year is selected – and showing the current and previous semester results, as seen on Figure 29, A. The absolute number of users that did complete the training and passed the test metric is shown on a KPI in figure 29, B. On detail C, it is shown the "Users not completing training" metric in two views, with relative and absolute figures, as previously noted. Several Year over Year comparison metrics in bar charts for the percentage and absolute figures of user's completeness are shown in Figure 29, D. Also shown is the number of users not attending to a training over category dimension, that is, by showing how many users from C Suite or management or regular category are not attending the mandatory trainings, this metric is demonstrated on Figure 29, E. It goes a step further in providing information regarding the average scoring of the training exam by country, as shown on Figure 29, detail G, where the colour gradient indicates the scoring, hence, the darker it is, the higher the score.

Lastly, a detail table regarding the training and exam statistics is provided for further analysis, in Figure 29, F. With Qlik associative engine, whenever the user clicks on any dimension in any object, it will assume it as a filter, therefore the data shown in each object/KPI will be filtered by that selection. For example, if a user selects Brazil in the map, all the metrics will provide the information directed to that country. However, four fields filters were provided in the left side of the sheet, the year, quarter, employee category and country. When selecting a quarter, no changes will be applied in the quarterly KPI's, it means that even though a user can select quarter 3, the KPI for quarter 1 will still display information regarding the correct quarter date as this was taken into account on defining the set analysis expression.

6.4.2. Password Quality

Although the information provided by this dashboard is related to metrics defined in 4.2.1, a new sheet on the application was created to show figures related with the password quality topic. On Figure 30, several implemented KPI's are provided, as for example on detail A, where the total number of passwords is shown, detail B provides the total number of complex passwords. Detail C shows the current ratio of compliance and detail D portraits a gauge showing the current password compliance status where a threshold is displayed. On detail E, a monthly analysis in a bar chart is provided with a trend analysis metric that provides some insight into future trends. Detail F provides a password detail table that give the opportunity for further insights into the current state of password quality. Finally, object depicted in detail G, helps the user obtain a fast analysis at a glance with the use of colour coding and alert icons. Whenever a KPI is above a certain defined threshold, the colour is green and red if the KPI did not reach the threshold, as shown in Figure 31, details B and C.

6.4.3. Social Engineering Metrics

The dashboard depicted at Figure 31 display information regarding a fictious phishing email sent to the employees of the organization and provides the latest assessment information. On A, the total number of emails sent is presented, B shows the percentage of users that followed the link on the phishing email, C shows the percentage of users which reported the email to the infosec team and D the number of social engineering incidents. Colour coding of the KPI's were used were appropriate to provide visual information whether a certain indicator as reached the threshold.

6.4.4. Incident Handling

The Incident handling dashboard shown on Figure 32 provides four general KPI's. The first one, Figure 32 A, is the cost of information incidents and it was calculated based on the number of hours to solve an incident multiplied by an average hourly wage[12]. The second, B, is the number of incidents in the current year and on the previous year (PY), C represents the number of actions for security improvements and finally the number of incidents not solved both on the current and previous year are shown on figure 32, D. Two charts are provided, the first one, E, analyses the number of incidents and the improvements over a month dimension and the second chart, depicted as F, displays monthly analysis of the incidents by category.

6.4.5. Vulnerabilities

The vulnerabilities dashboard, demonstrated as Figure 33, provides the visualization for the metrics previously defined in Section 4.2.3. It was extended to six more indicators to provide a wider overview of the situational awareness of the topic. As such, at this sheet the metrics presented are: Detail A, the number of assets and critical assets, object in detail B is the percentage of critical systems where vulnerability and *pentest* assessment has been done on detail C, the percentage of unpatched vulnerabilities is presented and the object in D is the ratio of organizational vulnerabilities coverage. Plus, detail E shows the percentage of *pentest* assessments made to all assets, detail F shows a gauge portraying the completeness of *pentest* ratio, object in detail G is the number of *pentest* assessments made to BIA[13] level 3 assets and on detail H, it is a Year to Year comparison of the number of *pentest* performed on a bar chart. The object group in detail I, shows three bar chart analysis by BIA level for assessment and patch coverage and the Year over Year by quarter for the number of assessments.

[12] Value obtained from https://yalantis.com/blog/cost-services-europe-market-research/

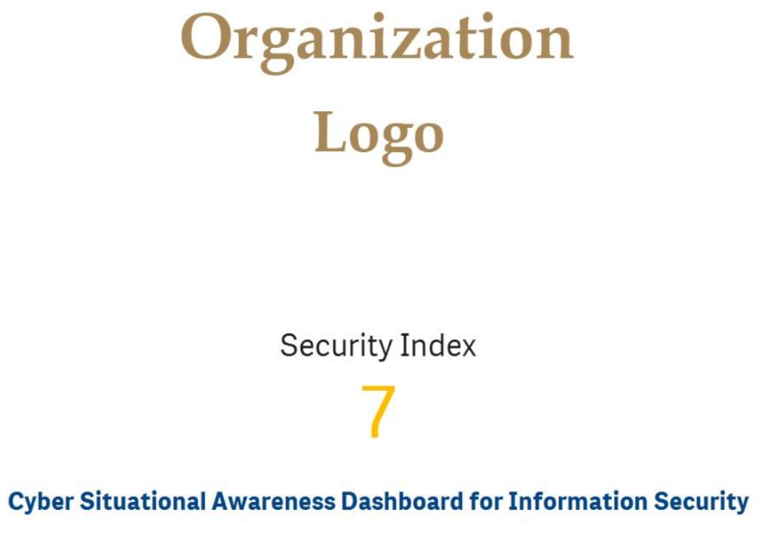
[13] BIA level 3 was assumed as the critical score in the context of this work.

6.4.6. Assets Inventory

The last section of this Situational Awareness Dashboard is the assets analysis of the organization. As is illustrated in Figure 34, it is shown the information regarding: A the number of authorized devices, B unauthorized devices, C the number of critical assets and D, the percentage of critical assets protected or patched. Figure E, shows a detail table that provides information regarding the authorized devices with the universally unique identifier (UUID), *IP* Address, System Name and business impact level of each assets. On section F, a table with the details about the unauthorized devices, such as the IP address and the system name, is also provided. Finally, on G section a comparison of Authorized vs Unauthorized is provided in a horizontal bar chart.

6.4.7. Assets Inventory

Lastly, security index displayed on the first page of the application provides the user with an overall overview about the situational security environment of the organization, as depicted in Figure 28. This metric was calculated for 6 randomly selected metrics in accordance to equation 4.5.1



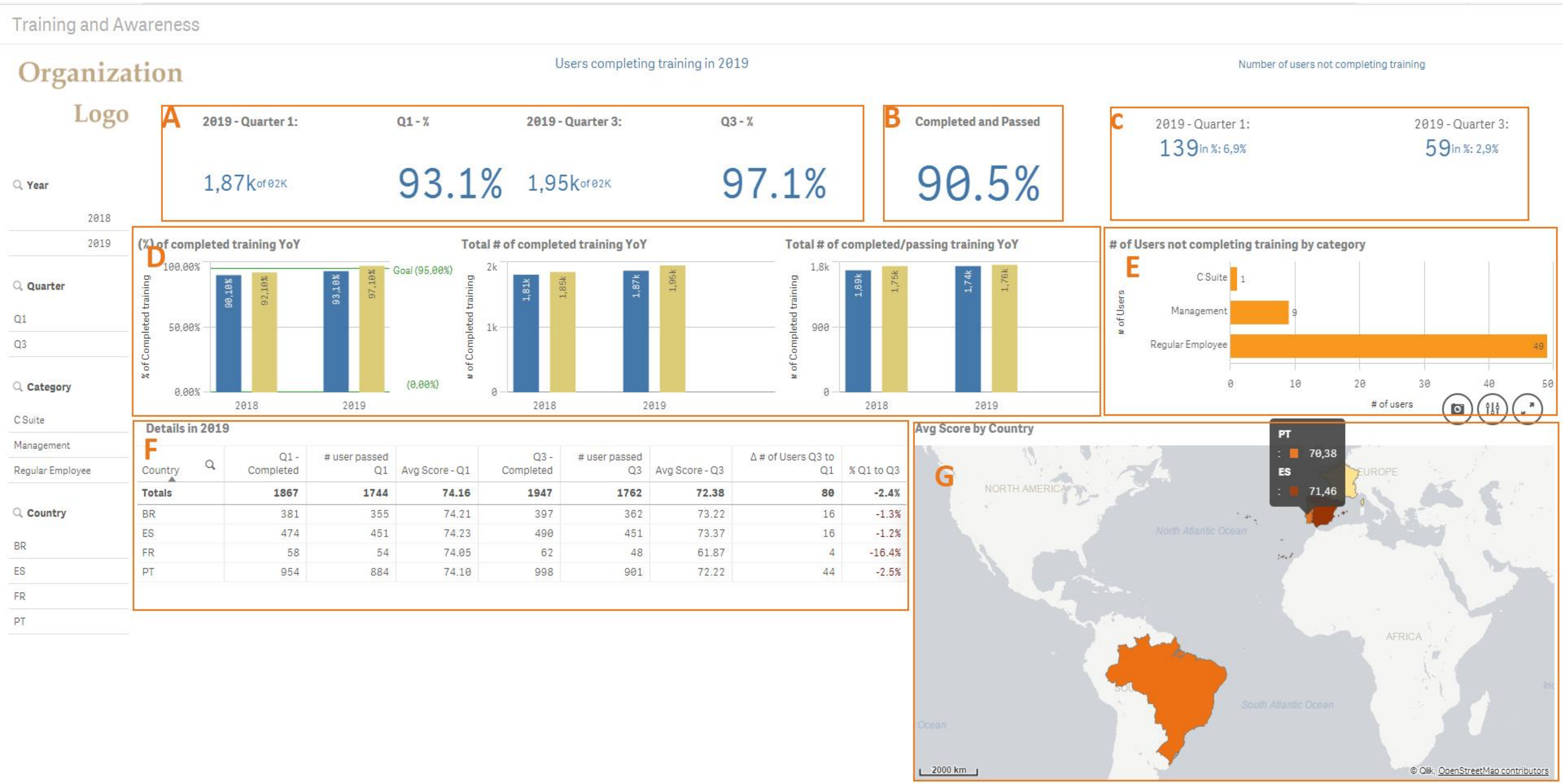
Organization Logo

Security Index

7

Cyber Situational Awareness Dashboard for Information Security

Figure 28 - Security Index



Training and Awareness

Users completing training in 2019

A 2019 - Quarter 1: 1,87k of 02K

Q1 - % 93.1%

2019 - Quarter 3: 1,95k of 02K

Q3 - % 97.1%

B Completed and Passed 90.5%

Number of users not completing training

C 2019 - Quarter 1: 139 in %: 6,9%

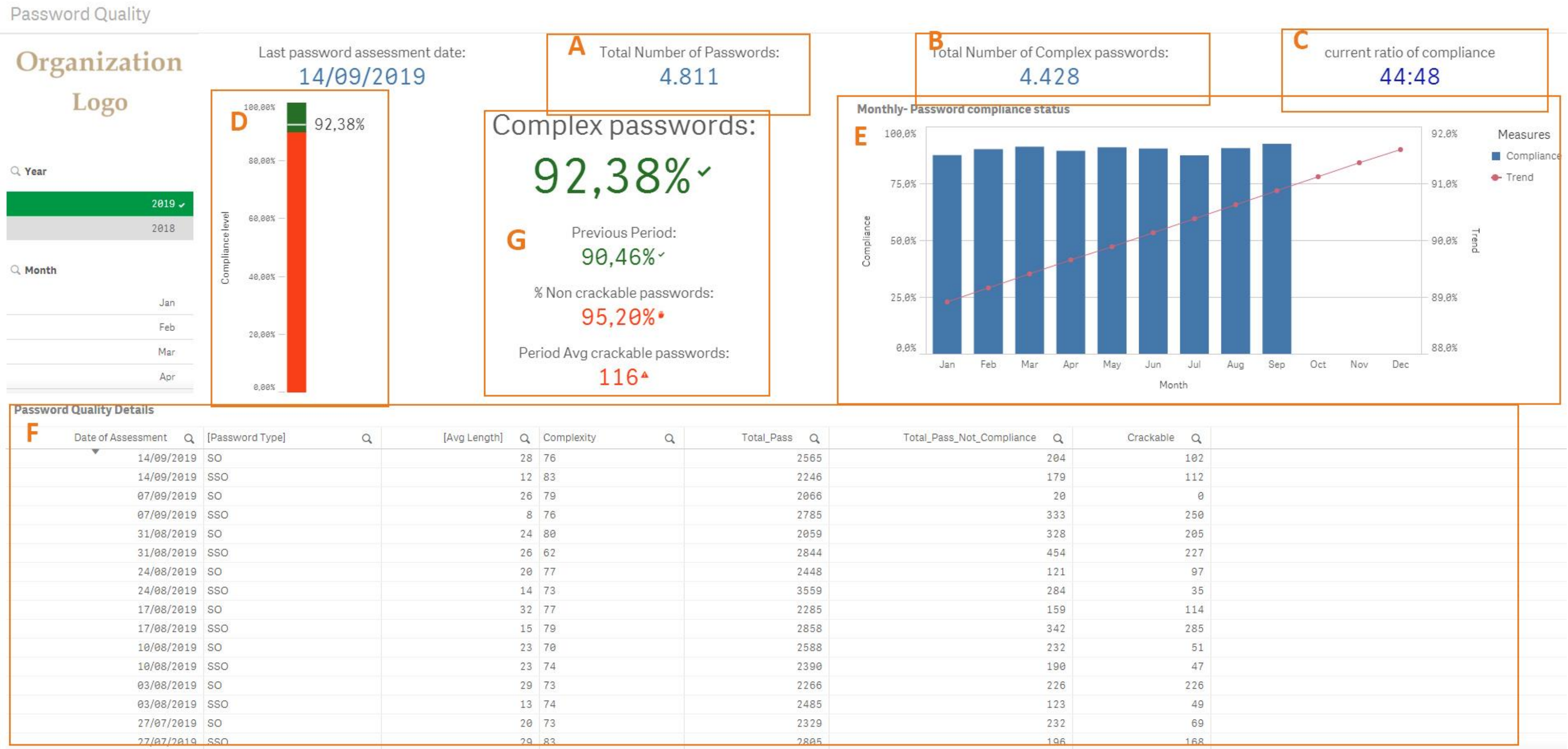
2019 - Quarter 3: 59 in %: 2,9%

Details in 2019

F

Country	Q1 - Completed	# user passed Q1	Avg Score - Q1	Q3 - Completed	# user passed Q3	Avg Score - Q3	Δ # of Users Q3 to Q1	% Q1 to Q3
Totals	**1867**	**1744**	**74.16**	**1947**	**1762**	**72.38**	**80**	**-2.4%**
BR	381	355	74.21	397	362	73.22	16	-1.3%
ES	474	451	74.23	490	451	73.37	16	-1.2%
FR	58	54	74.05	62	48	61.87	4	-16.4%
PT	954	884	74.10	998	901	72.22	44	-2.5%

Figure 29 - Training and Awareness Dashboard



Last password assessment date: 14/09/2019

Total Number of Passwords: 4.811

Total Number of Complex passwords: 4.428

current ratio of compliance 44:48

Complex passwords: 92,38% ✓

Previous Period: 90,46% ✓

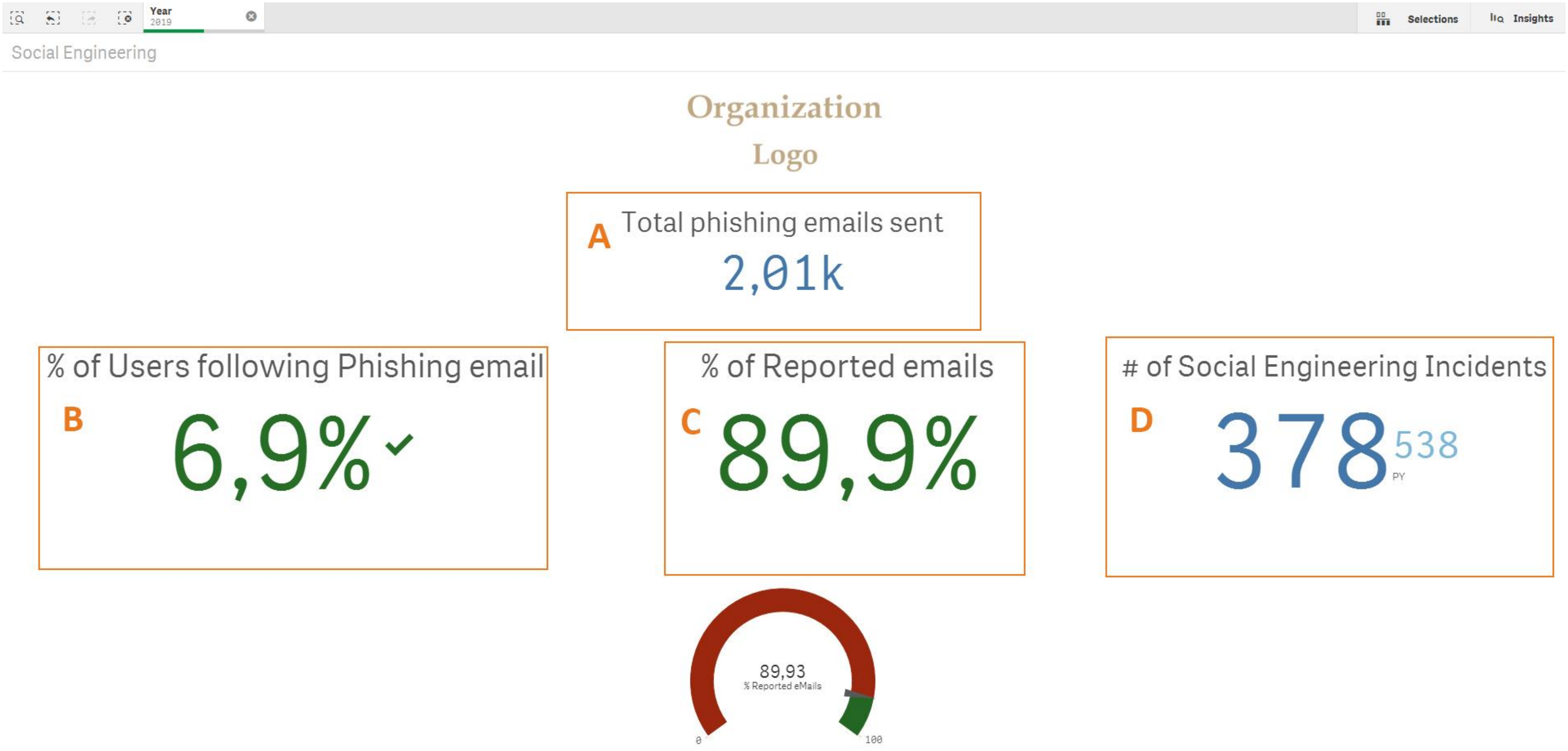
% Non crackable passwords: 95,20% ●

Period Avg crackable passwords: 116 ▲

Password Quality Details

Date of Assessment	[Password Type]	[Avg Length]	Complexity	Total_Pass	Total_Pass_Not_Compliance	Crackable
14/09/2019	SO	28	76	2565	204	102
14/09/2019	SSO	12	83	2246	179	112
07/09/2019	SO	26	79	2066	20	0
07/09/2019	SSO	8	76	2785	333	250
31/08/2019	SO	24	80	2059	328	205
31/08/2019	SSO	26	62	2844	454	227
24/08/2019	SO	20	77	2448	121	97
24/08/2019	SSO	14	73	3559	284	35
17/08/2019	SO	32	77	2285	159	114
17/08/2019	SSO	15	79	2858	342	285
10/08/2019	SO	23	70	2588	232	51
10/08/2019	SSO	23	74	2390	190	47
03/08/2019	SO	29	73	2266	226	226
03/08/2019	SSO	13	74	2485	123	49
27/07/2019	SO	20	73	2329	232	69
27/07/2019	SSO	29	83	2805	196	168

Figure 30 - Password Quality Dashboard



Year 2019

Selections

Insights

Social Engineering

Organization Logo

A Total phishing emails sent 2,01k

% of Users following Phishing email

B 6,9% ✓

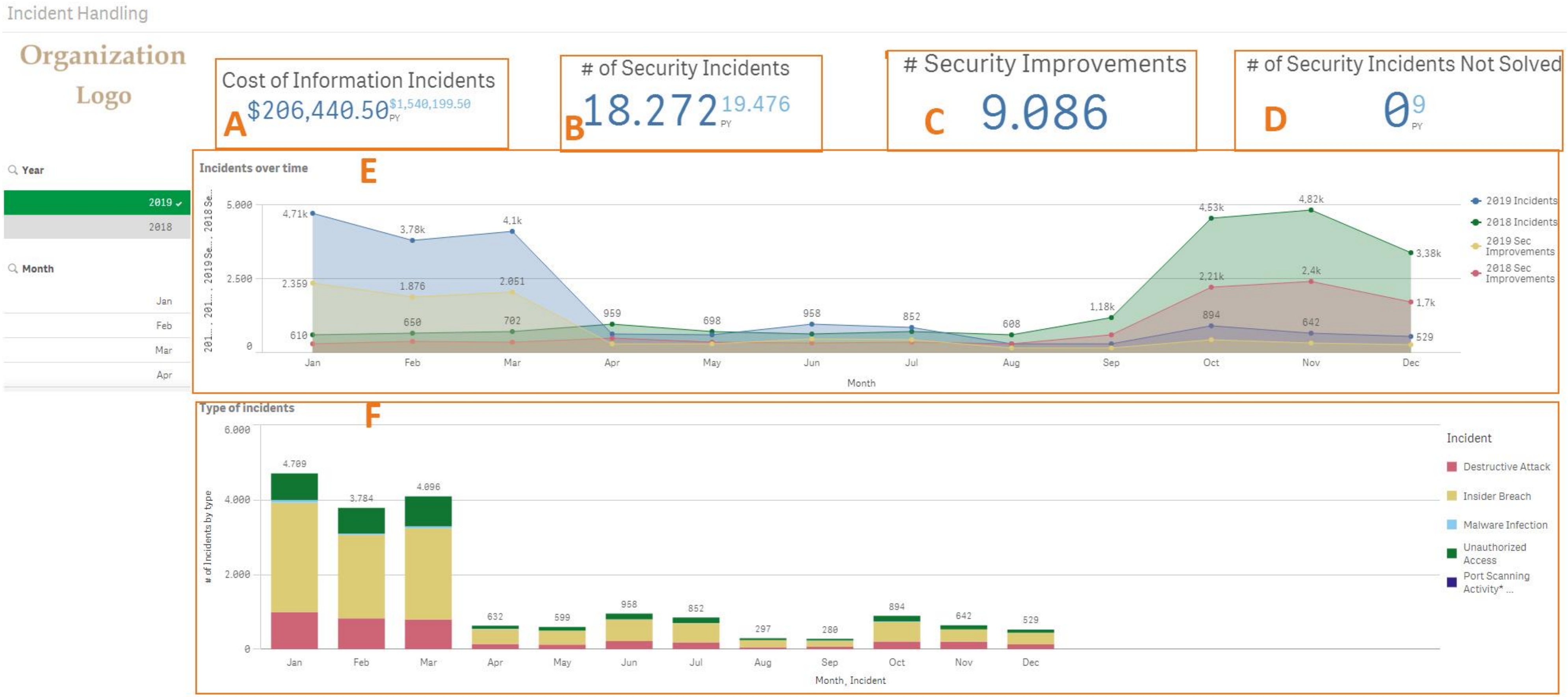
% of Reported emails

C 89,9%

# of Social Engineering Incidents

D 378 538 PY

Figure 31 - Social Engineering Dashboard



Incident Handling

Organization Logo

Cost of Information Incidents
$206,440.50 $1,540,199.50 PY

# of Security Incidents
18.272 19.476 PY

Security Improvements
9.086

# of Security Incidents Not Solved
0 9 PY

Year
2019 ✓
2018

Month
Jan
Feb
Mar
Apr

Figure 32 - Incident Handling Dashboard

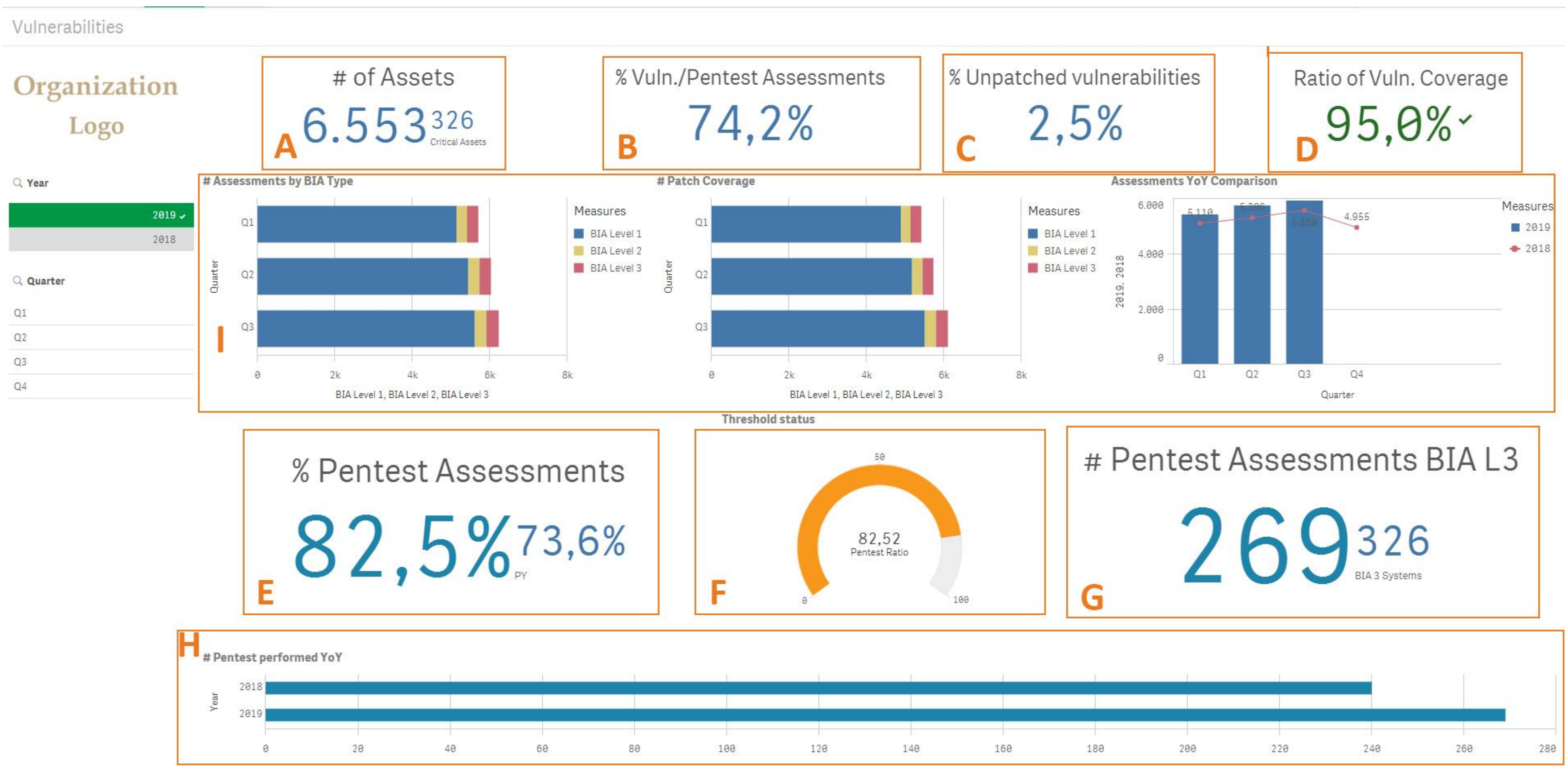
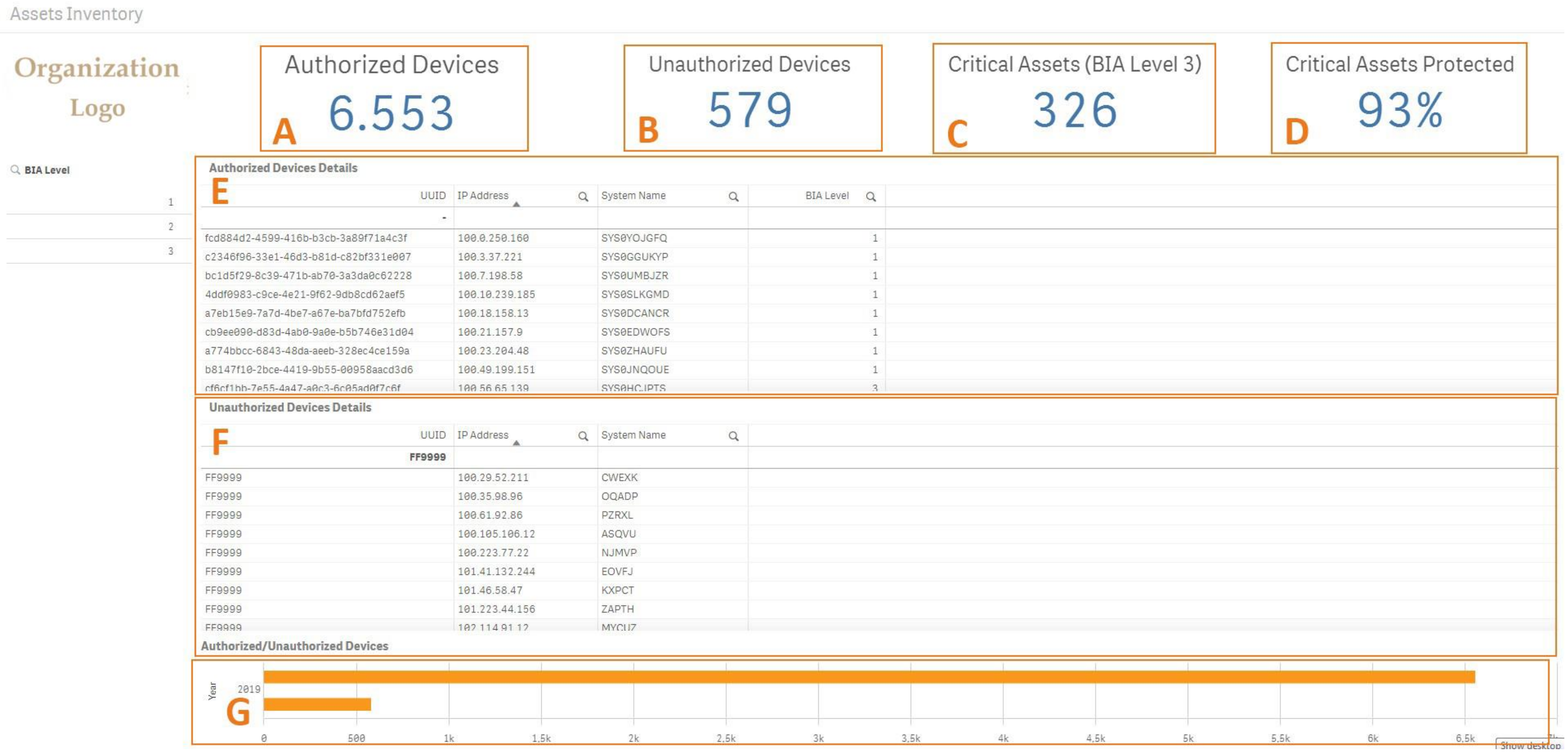


Figure 33 - Vulnerabilities Dashboard



Assets Inventory

Organization Logo

BIA Level

1

2

3

Authorized Devices

A 6.553

Unauthorized Devices

B 579

Critical Assets (BIA Level 3)

C 326

Critical Assets Protected

D 93%

Authorized Devices Details

E

UUID	IP Address	System Name	BIA Level
-			
fcd884d2-4599-416b-b3cb-3a89f71a4c3f	100.0.250.160	SYS0YOJGFQ	1
c2346f96-33e1-46d3-b81d-c82bf331e007	100.3.37.221	SYS0GGUKYP	1
bc1d5f29-8c39-471b-ab70-3a3da0c62228	100.7.198.58	SYS0UMBJZR	1
4ddf0983-c9ce-4e21-9f62-9db8cd62aef5	100.10.239.185	SYS0SLKGMD	1
a7eb15e9-7a7d-4be7-a67e-ba7bfd752efb	100.18.158.13	SYS0DCANCR	1
cb9ee090-d83d-4ab0-9a0e-b5b746e31d04	100.21.157.9	SYS0EDWOFS	1
a774bbcc-6843-48da-aeeb-328ec4ce159a	100.23.204.48	SYS0ZHAUFU	1
b8147f10-2bce-4419-9b55-00958aacd3d6	100.49.199.151	SYS0JNQOUE	1
cf6cf1bb-7e55-4a47-a0c3-6c05ad0f7c6f	100.56.65.139	SYS0HCJPTS	3

Unauthorized Devices Details

F

UUID	IP Address	System Name
FF9999		
FF9999	100.29.52.211	CWEXK
FF9999	100.35.98.96	OQADP
FF9999	100.61.92.86	PZRXL
FF9999	100.105.106.12	ASQVU
FF9999	100.223.77.22	NJMVP
FF9999	101.41.132.244	EOVFJ
FF9999	101.46.58.47	KXPCT
FF9999	101.223.44.156	ZAPTH
FF9999	102.114.91.12	MYCUZ

Authorized/Unauthorized Devices

Figure 34 - Assets Inventory Dashboard

6.5. Dashboard use cases

This section provides an overview about the possible uses of a Situational awareness dashboard. By using a dashboard, an organization keeps an up to date snapshot of the organization's information security risk management and ensure the visibility of the assets, therefore allowing the prioritization of actions, such as mitigation, acceptance or remediation if needed.

The first example portrays an example of use to evaluate the status of the information security training, how has it evolved over time and how it correlates with other overviews across the dashboard as for example, the phishing email testing.

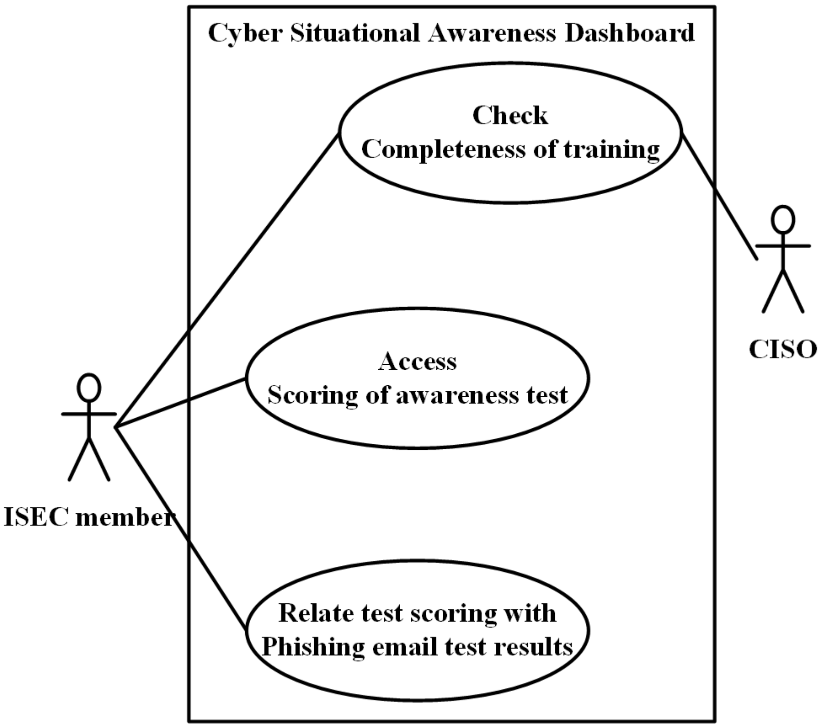

Cyber Situational Awareness Dashboard
Check
Completeness of training
Access
Scoring of awareness test
Relate test scoring with
Phishing email test results
ISEC member
CISO

Figure 35 - Dashboard use case of training effectiveness

On the Figure 35 an ISEC member responsible for the implementation of information security awareness program checks the outcomes of security training. The evolution across time can assist in the development of future trainings. It also provides a regional overview of the effectiveness of the training, therefore allowing to adapt the training to cultural differences. At the same time, a CISO can check if the threshold of training completeness and exam success is being achieved.

As an example, in Figure 29 Detail D, it can be seen that there is an increase in the training completeness over time. This type of analysis is based on comparison of one category, and, as such, a bar chart was used as it provides a fast and clear comparison.

More often than not, detailed analysis is useful to accomplish several comparisons at once, to achieve that, the information is provided in a tabular method, which provides the ability to compare across different dimensions, for example by date and country, helping to take action where appropriate. For example, on Figure 29, Detail F, it can be seen that for country France, there is a decrease in users passing in the exam and also a decrease on its average score. This analysis shows that some actions should be taken on the training methods for that country's employees in order to promote security awareness.

In Figure 29, Detail G, it is visible that the Spanish employees are the ones getting the best average scoring on this topic. By using a world map with a heat map and a colour gradient, this information is instantaneously passed to the user.

On the plus side, the dashboard user could select one of the countries in the map and all the other KPI's in the dashboard would adjust to provide the information in accordance to that selection by using Qlik's associative engine.

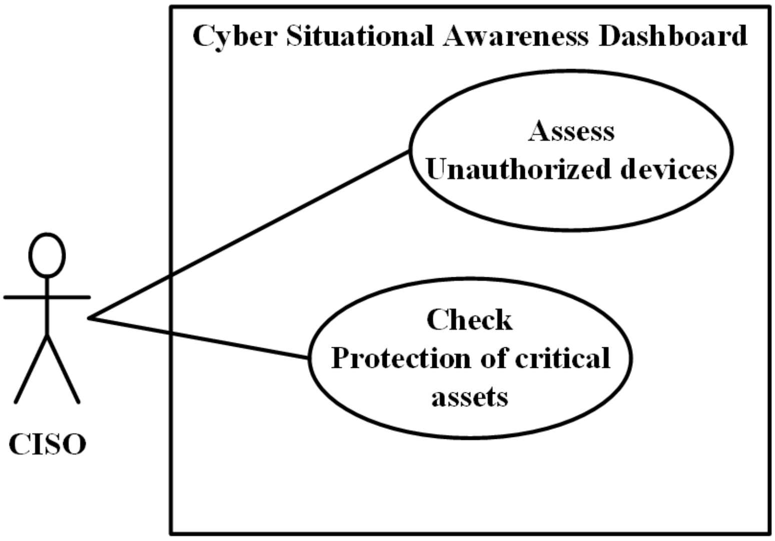

Cyber Situational Awareness Dashboard
Assess Unauthorized devices
Check Protection of critical assets
CISO

Figure 36 - CISO checks for unauthorized devices and the protection of critical assets

The second portrayed in Figure 36, example shows a CISO checking for unauthorized devices to be removed or inserted into the inventory and the percentage of protection of BIA level 3 assets. Combining this information with a threshold, the CISO has an instantaneous snapshot of the current situation that could trigger the necessary actions. In the example portrayed in Figure 34, Details C and D, the threshold of protection was reached, therefore no actions are needed on critical assets.

At the same time, Detail B of Figure 34, provides information regarding 579 unauthorized devices that are connected to the organization network. By assessing the details of these devices, as shown on Figure 34, Detail F, the CISO can act accordingly, by eliminating or updating such devices to the authorized devices list.

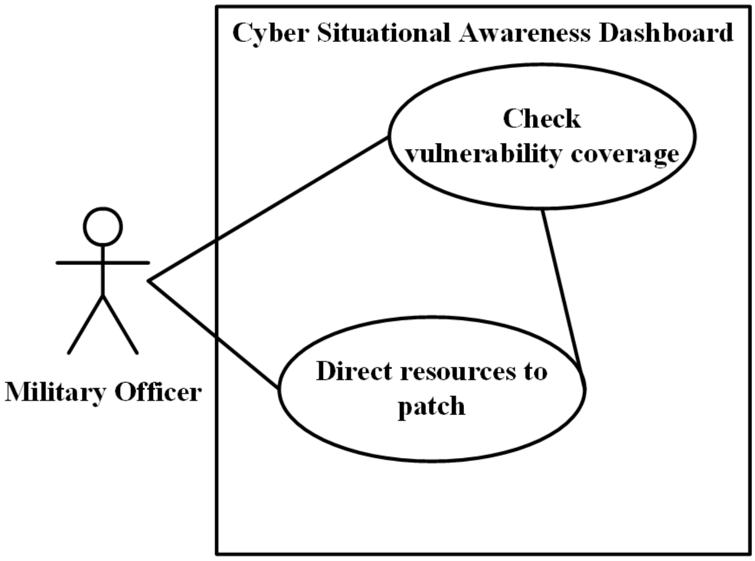

Cyber Situational Awareness Dashboard
Check vulnerability coverage
Direct resources to patch
Military Officer

Figure 37 - Military officer rate vulnerabilities in the network

The third example in Figure 37, depicts a military commander rating the vulnerability coverage of a network and redirecting resources to the patch process accordingly to the awareness obtained during the use of the dashboard. The example is obtained from Figure 33, Detail D and I, where it can be seen that the hypothetical threshold of vulnerability coverage was met and the number of patch coverage is depicted by asset category.

The fourth example in Figure 38, shows a CISO making an assessment of the number of incidents and the number of improvements triggered during the current and previous year. This analysis is done on a monthly dimension basis and provides an evolutionary overview of four metrics, providing a yearly comparison over the current and previous year. As such, a line chart was used, as it provides a clear visual analysis of the evolution of the metrics. It is clear that, less incidents triggers less security improvements and by this analysis it can be seen that both metric’s value has been declining, this correlation is a probable indicator that the security program is getting more efficient across time in the organization.

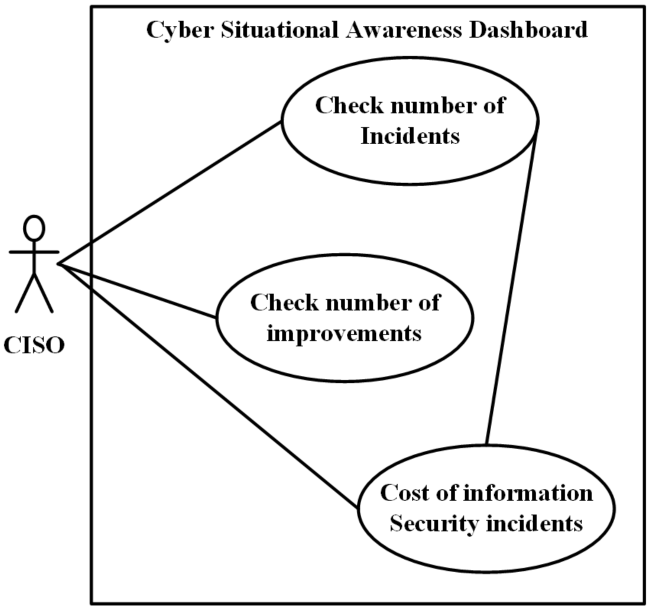

Cyber Situational Awareness Dashboard
Check number of Incidents
Check number of improvements
Cost of information Security incidents
CISO

Figure 38 - CISO create assessment of incidents and improvements

In the use case, shown on Figure 39, a CISO analyses the password compliance of the IT system users and check the prediction trend analysis for the next months, as illustrated on the combo chart in Figure 30, Detail E. By using this application, a user can also accomplish a self-service BI analysis approach, since the possibility of adding new visual objects to the analysis is allowed.

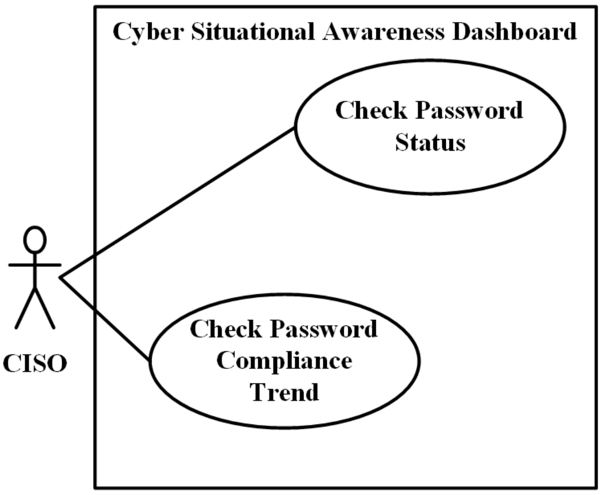

Cyber Situational Awareness Dashboard
Check Password Status
Check Password Compliance Trend
CISO

Figure 39 - CISO analyses password compliance

Appendix C provides a complete and clean overview of all the Dashboard sheets and visualizations implemented in the application.

Chapter 7
Conclusion

7.1. Conclusions

This dissertation aimed at creating a Situational Awareness Dashboard in order to provide information about cybersecurity to support decision takers in managing risk. To accomplish that goal, four specific objectives were initially defined in order to support the application development.

The first objective, which was to study information security metrics was accomplished and presented in section 2.4. In this section, a general study of the metrics concept was taken in which the definition of metrics and measures was introduced, the principles for effective measures and a methodology to guide the development of a metrics program was presented. Following this goal's requirement, a more exhaustive study of frameworks for information security metrics was executed and the results were presented in section 4.2 and 4.3. During the literature review, it was observed that the majority of the authors refer to *de facto* standards and frameworks, such as ISO/IEC 27000 series and the NIST 800, while other authors also mention several times the CVSS for vulnerabilities scoring and CIS controls for general security metrics. While the formers rely on organizational entities, the last two have their development based on security community efforts. The second objective was the collection and definition of security metrics in order to implement the CSADIS. Based on the first objective, a collection of metrics was directly extracted from multiple sources and is presented in Appendix A. This metric study evidenced the difficulty in defining metrics in an abstract context. It was shown that the process of selecting metrics is always dependent on the reality of an organization's requirements and it is an ongoing process in which the continuous improvement ultimately leads to a wide, efficient and mature model. Hence, some metrics were selected and adapted to implement in the tool developed in this work. The metrics definition is presented in section 4.4 and constitutes the base for the development of the CSADIS application.

The development of processes and architecture to sustain the implementation of the dashboard was the third objective. The study on the architecture was conducted and presented on Chapter 3 where SA architectures found on the literature were reviewed and the proposed functional architecture for the application was presented. This proposal was partially met, as the original architecture was intended for Qlikview, however, due to license constraints, only Qlik Sense license was made available. Thus, with the differences in the software, the architecture was adapted in order to meet the specificities of Qlik Sense. Chapter 5 materializes the previous study and presents the design and implementation of the Dashboard application. The difficulty in finding public security data to sustain the application led to the creation of several datasets. The creation of the datasets is presented in section 4.6 and illustrates all the steps taken to create useful data. Whenever possible, randomness was introduced in the creation of the datasets in order to create data similar to what would be expected from real-life systems. Since these datasets proved to be useful in this implementation, they will be available to the academic community in order to be used as desired for other works. Chapter 5 describes the development, design and implementation of the product of this work, the CSADIS application. Several steps are described, from the extraction of data from the sources, the construction of the data model, to the design of the frontend and creation of the metrics in set analysis expressions. Chapter 6 presented the Dashboard general overview with some use cases described, which demonstrates that the aim of this work was fully met.

A set of hypotheses leading the investigation of this work were initially defined. The first hypothesis defined that a regular BI tool would allow the construction and implementation of SA Dashboard instead of relying solely on commercial off-the-shelf (COTS) tool. The implementation of the CSADIS application demonstrates that this conception is true and that it is possible to create an application using such tools. The use of analytic and business intelligence tools provides means to collect data from diverse data sources, combining and presenting information using visual analytics, all developed through a tailor-made process that meets specific organization’s goals. Such type of dashboard is best suited for decision makers at a top management and executive-levels, i.e., CISO, CFO, ISEC team members and other relevant personnel in civil organization context. However, it also increases intelligence capabilities to military personnel so that strategic and tactical decisions can be taken in useful time or as a key component of information operations to centralize information and help planning actions. On the contrary, operational personnel usually rely on other type of analysis provided by commercial off-the-shelf (COTS) applications, typically at the transactional or operational levels. The use of a dashboard leverages both the operational and transactional data, by collecting, transforming and aggregating to an upper level and thus, exhibiting relevant information to the management level and senior executives from across an organization landscape.

The second hypothesis questioned whether a subset of security metrics would enable a value-drive product. Several distinct topics can be centralized in an application like CSADIS and the application modularity provides the means for further developments. This application provides relevant information and it is able to evolve accordingly to an organization's needs, therefore it enables value from the start of the implementation.

The last hypothesis defined that a CSADIS is a key component in information superiority. Based on the outcome of this study and the possibilities that an application such as the one developed on this work opens in terms of availability of information and possible integration of, not only predictive data, but also multiple general data sources, it is safe to assume that a CSADIS can be a key component in real life information systems.

This study proposed the implementation of an application capable of managing cyber security risk and, thus, generating value to an organization through the centralization of information from multiple data sources. Based on the above conclusions, it is safe to assume this work fulfilled its assumptions and objectives, by creating and demonstrating the value of a Cyber Situational Awareness Dashboard for Information Security.

7.2. Future work

Testing the Situational Awareness Dashboard developed in this work on an operational environment, with real large datasets and customized to the organization's reality, is one of the futures aim of this work. This would allow to have a final validation from end users of such systems, narrow down the focus on specific metrics and would open the possibility to test the use cases described in this work in order to evaluate the outcomes provided to an organization.

Another possibility for further developments in this work would be the use of predictive analytics to help manage the organization's risk. Using a dashboard like the outcome of this work opens up the possibility to connect to multiple sources, such as predictive software for data mining, deep learning and statistical analysis, therefore allowing the presentation of centralized information for further analysis and more efficient information security and risk management.

Reference List

Alberts, D. S., Garstka, J. J., Hayes, R. E., & Signori, D. T. (2001). *Understanding Information Age Warfare* (DoD CCRP (ed.)). DoD CCRP.

Alberts, D. S., Garstka, J. J., & Stein, F. P. (2000). *Network centric warfare: developing and leveraging information superiority* (DoD CCRP (ed.); Issue February). DoD CCRP.

Barabanov, R., Kowalski, S., & Yngstrom, L. (2019). Information security metrics: Research directions. *Controlled Information Security (COINS) Research Project*, 1–16.

Black, P. E., Scarfone, K. A., & Souppaya, M. P. (2008). Cyber Security Metrics and Measures. In *National Institute of Standards and Technology, National Institute of Standards and Technology*.

Borges, J. (2015). *Informação , Incerteza e Risco*. AcademiaMilitar, Lisboa, Portugal.

Bowen, P., Hash, J., & Wilson, M. (2006). Information Security Handbook: A Guide for Managers NIST Special Publication 800-100. In *National Institute of Standards and Technology Special Publication 800-100* (Issue October). National Institute of Standards and Technology.

Brooks, P., El-Gayar, O., & Sarnikar, S. (2015). A framework for developing a domain specific business intelligence maturity model: Application to healthcare. *International Journal of Information Management*, *35*(3), 337–345.

CCDCOE. (2012). *National Cyber Security Framework Manual* (N. C. C. Publications (ed.)). NATO.

Center for Internet Security. (2018). CIS Controls Measures and Metrics for Version 7. *CIS Website*.

Controls, C. I. S. (2018). *CIS Controls Basic Foundational Organizational*.

Dempsey, K., Chawla, N. S., Johnson, A., Johnston, R., Jones, A. C., Orebaugh, A., Scholl, M., & Stine, K. (2011). Information Security Continuous Monitoring (ISCM) for Federal Information Systems and Organizations. In *National Institute of Standards and Technology Special Publication 800-137*. National Institute of Standards and Technology.

Elizabeth Chew, Swanson, M., Stine, K., Bartol, N., Brown, A., & Robinson, W. (2008). Performance Measurement Guide for Information Security. In *National Institute of*

Standards and Technology Special Publication 800-55 (Issue July). National Institute of Standards and Technology.

English, D. (2010). *Understanding QlikView ' s Associative Architecture*.

ENISA. (2014). Actionable Information for Security Incident Response. *ENISA*, *November*, 1–79.

European Comission. (2019). *Objectives – CS-AWARE*. Retrieved April 2, 2019, from https://cs-aware.eu/objectives/

European Comission. (2013). *Improving cyber security across the EU - Consilium*. *February*. http://www.consilium.europa.eu/en/policies/cyber-security/

European Union. (2016). DIRECTIVE (EU) 2016/1148 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union. *Official Journal of the European Union*, *194*(March 2014), 1–30.

European Union. (2019). *Regulation (EU) 2019/881 of the European Parliament and of the Council on ENISA and on information and communications technology cybersecurity certification and repealing Regulation*. *2019*(March), 15–69.

European Union Agency for Network and Information Security. (2015). Definition of Cybersecurity | Gaps and overlaps in standardisation. In *European Union Agency For Network And Information Security: Vol. v1.0* (Issue December).

European Union Agency for Network and Information Security. (2017). *ENISA overview of cybersecurity and related terminology*. *1*, 1–8.

Few, S. (2013). Information dashboard design : displaying data for at-a-glance monitoring. *Information Dashboard Design : Displaying Data for at-a-Glance Monitoring.*, 166.

FIRST. (2019a). *Common Vulnerability Scoring System version 3.1 Specification Document Revision 1*. 1–24. https://www.first.org/cvss/

FIRST. (2019b). *Common Vulnerability Scoring System version 3.1 Specification Document Revision 1*. 1–24.

Franke, U., & Brynielsson, J. (2014). Cyber situational awareness - A systematic review of the literature. *Computers and Security*, *46*, 18–31.

Freund, J., & Jones, J. (2015). Information Security Metrics. In *Measuring and Managing Information Risk*. Elsevier Ltd.

García, M., & Harmsen, B. (2017). *QlikView for Developers*. Packt Publishing Ltd.

Grossmann, W., & Rinderle-Ma, S. (2015). *Fundamentals of Business Intelligence* (M. J. Carey & S. Cery (eds.)). Springer Berlin Heidelberg.

Han, W., Tian, Z., Huang, Z., Zhong, L., & Jia, Y. (2019). System architecture and key technologies of network security situation awareness system YHSAS. *Computers, Materials and Continua*, *59*(1), 167–180.

Howson, A. C., Richardson, J., Sallam, R., & Kronz, A. (2019). Magic Quadrant for Analytics and Business Intelligence Platforms. *Gartner*, 1–60.

International Organization for Standardization/International Electrotechnical Commission. (2011). Information technology - Security techniques - Information security risk management. In *ISO/IEC 27005:2011* (Vol. 2011).

International Organization for Standardization/International Electrotechnical Commission. (2013). Information security management systems - Requirements. In *ISO/IEC 27001:2013* (Vol. 2013, Issue ISO/IEC 27001:2013).

International Organization for Standardization/International Electrotechnical Commission. (2016). Monitoring, measurement, analysis and evaluation. In *ISO/IEC 27004:2016* (Vol. 2016).

International Organization for Standardization/International Electrotechnical Commission. (2018). Information security management systems - Overview and vocabulary. In *ISO/IEC 27000:2018* (5th ed.).

International Organization for Standardization. (2018). Risk Management - Guidelines. In *ISO 31000:2018* (2nd ed.).

Jim McCarthy, Alexander, O., Edwards, S., Faatz, D., Peloquin, C., Symington, S., Thibault, A., Wiltberger, J., & Viani, K. (2017). Situational Awareness. In *National Institute of Standards and Technology Special Publication 1800-7* (p. 241). SAGE Publications, Inc.

Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling* (JOHN WILEY & SONS INC (ed.); Third Edit). John Wiley & Sons.

Korpela, K. (2015). Improving Cyber Security Awareness and Training Programs with Data Analytics. *Information Security Journal*, *24*(1–3), 72–77.

Kott, A., Wang, C., & Erbacher, R. F. (2014). Cyber Defense and Situational Awareness: Foundations and Challenges. In *Advances in Information Security* (Vol. 62). Springer.

Laudy, C., Mattioli, J., & Museux, N. (2006). Cognitive Situation Awareness for Information Superiority. *NATO Information Fusion for Command Support*, *Bsik 03024*, 3-1-3–12.

Liu, P., Jajodia, S., & Wang, C. (2017). *Theory and Models for Cyber Situation Awareness*. Springer.

Matthews, E. D., Arata III, H. J., & Hale, B. L. (2016). Cyber Situational Awareness. *The Cyber Defense Review*, *1*(1), 209–233.

Mell, Peter; Waltermire, David; Feldman, Larry; Booth, Harold; Ouyang, Alfred; Ragland, Zach; McBride, T. (2012). CAESARS Framework. Extension: An Enterprise Continuous Technical Reference Model. In *National Institute of Standards and Technology Interagency Report 7756* (Vol. 56). National Institute of Standards and Technology.

Michael Hoehl. (2010). Creating a monthly information security scorecard for CIO and CFO. *Methods for Understanding and Reducing Social Engineering Attacks*, 36.

National Institute of Standards and Technology. (2011). Managing Information Security Risk Organization, Mission, and Information System. In *NIST Special Publication 800-39* (Issue March). National Institute of Standards and Technology.

National Institute of Standards and Technology. (2012). Guide for Conducting Risk Assessments. In *NIST Special Publication 800-30* (Issue September). National Institute of Standards and Technology.

National Institute of Standards and Technology. (2018). Risk Management Framework for Information Systems and Organizations A System Life Cycle Approach for Security and Privacy. In *NIST Special Publication 800-37: Vol. Rev 2* (Issue December). National Institute of Standards and Technology.

Noel, S., & Heinbockel, W. (2015). An Overview of MITRE Cyber Situational Awareness Solutions. *Mitre Corporation*, *May*, 1–17.

Note, J. D., Commander, J. F., Head, S., & Conditions, S. (2013). Joint Doctrine Note 2 / 13 Information Superiority. *Development, Concepts and Doctrine Centre - Ministry of Defense*.

Nunes, P. V. (2015). *Sociedade em rede, ciberespaço e guerra de informação : contributos para o enquadramento e construção de uma estratégia nacional de informação* (I. da D. Nacional (ed.); 1st ed.). Atena.

Nunes, P. V., Mendes, C. P., Ralo, J., Santos, L., Santos, L. C. dos, Moniz, P., & Casimiro, S. de V. (2018). Contributos para uma Estratégia Nacional de Ciberdefesa. In Instituto da Defesa Nacional (Ed.), *Instituto da Defesa Nacional* (Issue 28).

Pareek, M. (2017). Standardized Scoring for Security and Risk Metrics. *ISACA Journal*, *2*, 1–6.

Payne, S. (2006). A Guide to Security Metrics. *SANS Institute Reading Room*.

PCI Security Standards Council. (2014). Best Practices for Implementing a Security

Awareness Program. *PCI Data Security Standard (PCI DSS), October,* 27.

Pescatore, J. (2017). Back to Basics : Focus on the First Six CIS Critical Security Controls. *SANS Institute Reading Room*.

Peters, M. D., Wieder, B., Sutton, S. G., & Wakefield, J. (2016). Business intelligence systems use in performance measurement capabilities: Implications for enhanced competitive advantage. *International Journal of Accounting Information Systems, 21*, 1–17.

Qlik. (2017). *The Associative Difference ™ Relational Databases and Queries are Old Technology*. 4.

Rathbun, D. (2009a). *Gathering security metrics and reaping the rewards*.

Rathbun, D. (2009b). Gathering Security Metrics and Reaping the Rewards. *SANS Institute Reading Room*, 21.

Russom, P., Halper, F., Stodder, D., & Berberich, M. (2010). *12 Characteristics of Effective Metrics | Transforming Data with Intelligence*. TDWI. https://tdwi.org/Blogs/TDWI-Blog/2010/04/Effective-Metrics.aspx

Tashi, I., & Ghernaouti-Hélie, S. (2007). Security metrics to improve information security management. *Proceedings of the 6th Annual Security Conferenceth Annual Security Conference*, 47–1--47–13.

Tianfield, H. (2016). Cyber Security Situational Awareness. *Proceedings - 2016 IEEE International Conference on Internet of Things; IEEE Green Computing and Communications; IEEE Cyber, Physical, and Social Computing; IEEE Smart Data, IThings-GreenCom-CPSCom-Smart Data 2016*, 782–787.

Tufte, E. R. (2007). The Visual Display Of Quantitative Information. In *The Graphic Press* (p. 191).

Voeller, J. G., Black, P. E., Scarfone, K., & Souppaya, M. (2008). Cyber Security Metrics and Measures. *Wiley Handbook of Science and Technology for Homeland Security*, 1–8.

Wang, C. H. (2016). A novel approach to conduct the importance-satisfaction analysis for acquiring typical user groups in business-intelligence systems. *Computers in Human Behavior, 54*, 673–681.

What is QlikView? (n.d.). Retrieved May 23, 2019, from https://help.qlik.com/en-US/qlikview/April2019/Content/QV_HelpSites/what-is.htm

8.

Appendix A – Summary of relevant metrics

This Appendix presents a direct copy of several metrics found in the literature. This is not an adaptation nor original work by the author.

Table 5 - Metrics

Metric	Measaure/Description	Source
Asset capacity	The (remained) capacity of a cyber asset (after being attacked or compromised)	**(Kott, Wang and Erbacher, 2014)**
Average length of attack paths	The average effort to penetrate a network, or compromise a system/service; evaluated by attack graphs	
Compromised host percentage	The percentage of compromised hosts in a network at time *t*	
Exploit probability	How easy (or hard) to exploit a vulnerability? Could be measured by CVSS exploitability sub-score	
Impact factor	The impact level of a vulnerability after being exploited, could be measured by CVSS impact sub-score	
Number of attack paths	The number of potential attack paths in a network, could be evaluated based on attack graphs	
Network preparedness	Is a network ready to carry out a mission? E.g., all required services are supported by available cyber assets	
Network resilience	The percentage of compromised systems/services that can be replaced/ recovered by backup/alternative systems/services	
Operational capacity	The (remained) operational capacity of a system/service (after being affected by a direct attack or indirect impact)	
Resource redundancy	Is there any redundant (backup) resources assigned or allocated for a critical task/operation?	
Service availability	The availability of a required service to support a particular mission, task or operation	
Shortest attack path	The minimal effort to penetrate a network, or compromise a system or service, evaluated by attack graphs	
Severity score	The severity/risk of a vulnerability if it was successfully exploited, could be measured based on CVSS score	
Vulnerable host percentage	The percentage of vulnerable hosts in a network	
Security Budget	Percentage (%) of the agency's information system budget devoted to information	**(NIST Special Publication 800-55, 2008)**
Vulnerability	Percentage (%) of high vulnerabilities mitigated within organizationally defined time periods after discovery	
Remote Access Control Measure	Percentage (%) of remote access points used to gain unauthorized access	

Awareness and Training	Percentage (%) of information system security personnel that have received security training	
Audit Record Review	Average frequency of audit recods review and analysis for inappropriate activity	
Certification, Accreditation, and Security Assessments	Percentage (%) of new systems that have completed certification and accreditation (C&A) prior to their implementation	
Configuration Management	Percentage (%) approved and implemented configuration changes identified in the latest baseline configuration	
Contingency Planning	Percentage (%) of information systems that have conducted annual contingency plan testing	
Identification and Authentication	Percentage (%) of users with access to shared accounts	
Incident Response	Percentage (%) of incidents reported within required time frame per applicable incident category	
Maintenance	Percentage (%) of systems components that undergo maintenance in accordance with formal maintenance schedules	
Media Protection	Percentage (%) of media that passes sanitization procedures testing for FIPS 199 high-impact systems	
Physical and Environment	Percentage (%) of physical security incidents allowing unauthorized entry into facilities containing information systems	
Planning	Percentage of employees who are authorized access to information systems only after they sign an acknowledgement that they have read and understood rules of behaviour	
Personnel Security	Percentage (%) of individuals screened before being granted access to organizational information and information systems	
Risk Assessment	Percentage (%) of vulnerabilities remediated within organization-specified time frames	
System and Services Acquisition	Percentage (%) of system and service acquisition contracts that include security requirements and/or specifications	
System and Communications Protection	Percentage of mobile computers and devices that perform all cryptographic operations using FIPS 140-2 validated cryptographic modules operating in approved modes of operation	
System and Information Integrity	Percentage (%) of operating system vulnerabilities for which patches have been applied or that have been otherwise mitigated	
Resource allocation	Breakdown of resources allocated to information security (internal personnel, contracted personnel, hardware, software, services) within annual budget	**(ISO/IEC 27004:2016, 2016)**
Policy review	Percentage of policy reviewed	
Management commitment	a) Management review meetings completed to date b) Average participation rates in management review meetings to date	
Risk Exposure	a) High and medium risks beyond acceptable threshold b) Timely review of high and medium risks	
Audit programme	Total number of audits performed compared with the total number of audits planned	
Improvement actions	Percentage of actions on time, costs and quality (i.e. requirements) against all planned actions The actions should be the ones planned (i.e. opened, stand-by and in progress) in the beginning of the timeframe	
Security incident cost	Sum of costs for each information security incident occurred in the sampling period	
Learning from information security incidents	Number of security incidents that trigger information security improvement actions	
Corrective action implementation	a) Status expressed as a ratio of corrective action not implemented b) Status expressed as a ratio of corrective action not implemented without reason c) Trend of statuses	
ISMS training or ISMS awareness	Percentage of employees having participated to an ISMS awareness training	
Information security training	Percentage of personnel who received annual information security awareness training	

Information security awareness compliance	1. Progress to date 2. Progress to date with signing	
ISMS awareness campaigns effectiveness	Percentage of employees passing a knowledge test before and after ISMS awareness campaign	
Social engineering preparedness	Percentage of staff that react correctly to a test, e. g., who did not click on a link in a given test consisting in sending a phishing email to (a selected part of the) staff	
Password quality – manual	Total number of passwords that comply with organization’s password quality policy a) Ratio of passwords which meet organization’s password quality policy b) Trends of compliance status regarding password quality policy	
Password quality – automated	1 Total number of passwords 2 Total number of uncrackable passwords	
Review of user access rights	Percentage of critical systems where user access rights are periodically reviewed	
Physical entry controls system evaluation	Strength of physical entry controls system	
Physical entry controls effectiveness	Number of unauthorized entry into facilities containing information systems (subset of physical security incidents)	
Management of periodic maintenance	For each completed event, subtract [Date of actual maintenance] from [Date of scheduled maintenance]	
Change management	Percentage of new installed systems that were respected change management best practice and hardening policy	
Protection against malicious code	Trend of detected attacks that were not blocked over multiple reporting periods	
Anti-malware	Percentage of malware affected systems connected to the organization’s network with obsolete (e.g. more than one week) antimalware signatures	
Total availability	For each IT service the end-to-end availability is compared with the maximum availability (i.e., excluding the previously defined downtime windows)	
Firewall rules	Unused firewall rules on border firewalls	
Log files review	Percentage of audit log files reviewed when required per time period	
Device configuration	Percentage of devices (by type) configured according to policy	
Pentest and vulnerability assessment	Percentage of critical information systems where a penetration test or vulnerability assessment has been executed since their last major release	
Vulnerability landscape	Weight of open (unpatched) vulnerabilities	
Security in third party agreements – A	Average percent of relevant security requirements addressed in third party agreements	
Security in third party agreements – B	Average percent of relevant security requirements addressed in third party agreements - Personal information processing	
Information security incident management effectiveness	Incidents not resolved in target timeframe	
Security incidents trend	1. Number of information security incidents in a defined timeframe (e.g., month) 2. Number of information security incidents of a specific category in a defined timeframe (e.g., month)	
Security event reporting	Sum of security events reported to the Computer security incident response team (CSIRT) in relation to the size of the organization	
ISMS review process	Progress ratio of accomplished independent reviews	
Vulnerability coverage	Ratio of systems which have been object of vulnerability assessment/penetration testing activities	
Phishing Awareness	Number of people who fall victim to a phishing simulation. The definition of falling victim is clicking on the link or opening an attachment.	**SANS Security Awareness Metrics**
Phishing Reporting	Number of people who detect and report a phishing email (regardless of whether it's an assessment or real attack).	

Phishing Repeat Offenders	Number of workforce that repeatedly fall victim to phishing simulations. These individuals are not changing behavior and represent a high risk.	
Facility Physical Security	Number of employees who understand, follow, and enforce your policies for restricted or protected access to facilities.	
Updated Devices	Percentage of devices that are updated and current.	
Lost/Stolen Devices	Number of devices (laptops, smartphones, tablets) that were lost or stolen. What percentage of those devices were encrypted?	
Secure Desktop	Number of employees who are securing their desk environment before leaving, as per organizational policy.	
Passwords	Number of employees using strong passwords.	
Social Engineering	Number of employees who can identify, stop, and report a social engineering attack.	
Sensitive Data	Number of employees posting sensitive organizational information on social networking sites.	
Data Wiping or Destruction	Number of employees who are properly following data destruction processes.	
Device Physical Security	Number of employees who left their devices unsecured in their cars in the organization's parking lot.	
Engagement	Number of requests the security awareness team gets to do security briefings for other business units or teams	
Knowledge	Does workforce know and understand what is expected of them?	
Workforce's attitudes towards security	Does the workforce understand the need for security, the important role they play, and support the behaviors needed?	
Time to Detect an Incident	What is the average time it takes to detect an incident?	
Policy Violations	Number of times workforce violates organizational security policies.	
Data Loss Incidents	Number of times there is a data loss incident, either accidental or due to a deliberate attack.	
Infected Computers	Number of infected computers.	
Privileged Account Abuse	Number of privileged users that improperly use or abuse their privileged access.	
Misconfigured Systems	Number of incidents of systems or applications misconfigured.	
Compliance or Audit Violations	Number of compliance or audit violations or fines.	
Training Completion	Who has or has not completed annual security awareness training	
Communication Methods	Types of reinforcement training, who is consuming that training, and how often	
Policy Sign-Off	Ensuring employees have completed training, acknowledge they understand the training, and will adhere to the policies	
Number of Ambassadors	Number of active Ambassadors promoting the security awareness program	
Number of people Ambassadors are reaching	Combine the total number of your workforce that all of the Ambassadors are reaching.	
Number of times workforce is engaging Ambassadors	How often the ambassadors are approached with security questions, requests to present locally, or other engagement types.	
Effectiveness of Ambassadors	Compare the security awareness impact metrics in departments or offices that do have Ambassadors vs. those that do not.	
Number and type of outreach activities by Ambassadors	How often does each ambassador engage or communicate to their local team, and what is the communication method?	

Success stories	Real-world stories on how workforce identified and/or stopped a real attack.	
Surveys	Workforce's attitudes, beliefs, and certain behaviors.	
Detection of Incidents	How many incidents were detected, how fast, and by whom / how?	
Average time spent to manage each Ambassador	How much time are you spending managing each Ambassador?	
Utilize an Active Discovery Tool	What percentage of the organization's networks have not recently been scanned by an active asset discovery tool?	**(Center for Internet Security, 2018)**
Use a Passive Asset Discovery Tool	What percentage of the organization's networks are not being monitored by a passive asset discovery tool?	
Use DHCP Logging to Update Asset Inventory	What percentage of the organization's DHCP servers do not have logging enabled?	
Maintain Detailed Asset Inventory	What percentage of the organization's hardware assets are not presently included in the organization's asset inventory?	
Maintain Asset Inventory Information	What percentage of the organization's hardware assets as a whole are not documented in the organization's asset inventory with the appropriate network address, hardware address, machine name, data asset owner, and department for each asset?	
Address Unauthorized Assets	What percentage of the organization's unauthorized assets have not been removed from the network, quarantined or added to the inventory in a timely manner?	
Deploy Port Level Access Control	What percentage of the organization's network switches are not configured to require network-based port level access control for all client connections?	
Utilize Client Certificates to Authenticate Hardware Assets	What percentage of the organization's network switches are not configured to require network-based port level access control utilizing client certificates to authenticate all client connections?	
Maintain Inventory of Authorized Software	What percentage of the organization's software are not presently included in the organization's software inventory?	
Ensure Software is Supported by Vendor	What percentage of the organization's software applications or operating systems are not currently supported by the software's vendor?	
Utilize Software Inventory Tools	What percentage of the organization's hardware assets have not recently been scanned by a software inventory tool to document the software installed on the system?	
Track Software Inventory Information	What percentage of software assets are not documented in a software inventory system that tracks the name, version, publisher, and install date for all software, including operating systems authorized by the organization?	
Integrate Software and Hardware Asset Inventories	Is the organization's software inventory system tied into the hardware asset inventory system?	
Address unapproved software	What percentage of the organization unauthorized software are either removed or the inventory is updated in a timely manner?	
Utilize Application Whitelisting	What percentage of the organization's hardware assets are not utilizing application whitelisting technology to block unauthorized applications from executing on the system?	
Implement Application Whitelisting of Libraries	What percentage of the organization's hardware assets are not utilizing application whitelisting technology to block unauthorized applications at the library level from executing on the system?	
Implement Application Whitelisting of Scripts	What percentage of the organization's hardware assets are not utilizing application whitelisting technology to block unauthorized scripts from executing on the system?	
Physically or Logically Segregate High Risk Applications	What percentage of high risk business applications have not been physically or logically segregated from other business systems?	
Run Automated Vulnerability Scanning Tools	What percentage of the organization's hardware assets have not recently been scanned by an SCAP compliant configuration monitoring system to identify all potential vulnerabilities on the organization's systems?	
Perform Authenticated Vulnerability Scanning	What percentage of the organization's hardware assets have not recently been scanned by an SCAP compliant configuration monitoring system to identify all potential vulnerabilities on the organization's systems utilizing an authenticated connection to the system?	
Protect Dedicated Assessment Accounts	What percentage of the organization's hardware assets have not recently been scanned by an SCAP compliant configuration monitoring system to identify all potential vulnerabilities on the organization's systems utilizing a dedicated service account and host-based restrictions?	
Deploy Automated Operating System Patch Management Tools	What percentage of the organization's hardware assets are not regularly updated by an automated software update tools in order to ensure that the operating systems are running the most recent security updates provided by the software vendor?	

Deploy Automated Software Patch Management Tools	What percentage of the organization's hardware assets are not regularly updated by an automated software update tools in order to ensure that third-party software is running the most recent security updates provided by the software vendor?	
Compare Back-to-back Vulnerability Scans	What percentage of the organization's identified vulnerabilities have not been remediated in a timely manner?	
Utilize a Risk-rating Process	Has the organization utilized a risk-rating process to prioritize the remediation of discovered vulnerabilities?	
Maintain Inventory of Administrative Accounts	What percentage of the organization's hardware assets have not recently utilized automated tools to inventory all administrative accounts to ensure that only authorized individuals have elevated privileges?	
Change Default Passwords	What percentage of the organization's systems utilize default passwords for accounts with elevated capabilities?	
Ensure the Use of Dedicated Administrative Accounts	What percentage of the organization's user accounts with elevated rights do not utilize a dedicated or secondary account for elevated activities?	
Use Unique Passwords	What percentage of the organization's systems, where multi-factor authentication is not supported (such as local administrator, root, or service accounts), accounts will use passwords that are unique to that system?	
Use Multifactor Authentication For All Administrative Access	What percentage of the organization's hardware assets are not configured to utilize multi-factor authentication and encrypted channels for all elevated account access?	
Use of Dedicated Machines For All Administrative Tasks	What percentage of the organization's system administrators are not required to use a dedicated machine for all administrative tasks or tasks requiring elevated access?	
Limit Access to Script Tools	What percentage of the organization's systems limit access to scripting tools (such as Microsoft PowerShell and Python) to only administrative or development users with the need to access those capabilities?	
Log and Alert on Changes to Administrative Group Membership	What percentage of the organizations hardware assets are not configured to issue a log entry and alert when an account is added to or removed from any group assigned elevated privileges?	
Log and Alert on Unsuccessful Administrative Account Login	What percentage of the organization's hardware assets are not configured to issue a log entry and alert on unsuccessful logins to an administrative account?	
Establish Secure Configurations	What percentage of the organization's authorized operating systems and software does not have a documented, standard security configuration?	
Maintain Secure Images	What percentage of the organization's hardware assets are not based upon secure images or templates based on the organization's approved configuration standards?	
Securely Store Master Images	What percentage of the organization's master images are not stored on securely configured servers, validated with integrity checking tools, to ensure that only authorized changes to the images are possible?	
Deploy System Configuration Management Tools	What percentage of the organization's hardware assets are not automatically configured via system configuration management tools that automatically enforce and redeploy configuration settings to systems at regularly scheduled intervals?	
Implement Automated Configuration Monitoring Systems	What percentage of the organization's hardware assets have not recently been scanned by an SCAP compliant configuration monitoring system to verify all security configuration elements, and alert when unauthorized changes occur?	
Utilize Three Synchronized Time Sources	What percentage of the organization's hardware assets do not utilize at least three synchronized time sources from which all servers and network devices retrieve time information on a regular basis so that timestamps in logs are consistent?	
Activate audit logging	What percentage of the organization's hardware assets are not configured to require local logging on the asset?	
Enable Detailed Logging	What percentage of the organization's hardware assets are not configured to require local logging to include detailed information such as a event source, date, timestamp, source addresses, destination addresses, and other useful elements on the asset?	
Ensure adequate storage for logs	What percentage of the organization's hardware assets do not have adequate storage space for the logs generated?	
Central Log Management	What percentage of the organization's hardware assets are not configured to aggregate appropriate logs to a central log management system for analysis and review?	
Deploy SIEM or Log Analytic tool	What percentage of the organization's hardware assets are not configured to aggregate appropriate logs to a Security Information and Event Management (SIEM) or log analytic tools for log correlation and analysis?	
Regularly Review Logs	What percentage of the organization's hardware assets have not had their logs reviewed recently to identify anomalies or abnormal events?	
Regularly Tune SIEM	What percentage of the organization's SIEM systems have not recently been tuned to better identify actionable events and decrease event noise?	
Ensure Use of Only Fully Supported Browsers and Email Clients	What percentage of the organization's hardware assets are running unsupported web browsers and email client software?	

Disable Unnecessary or Unauthorized Browser or Email Client Plugins	What percentage of the organization's hardware assets are utilizing unauthorized browser or email client plugins or add-on applications?	
Limit Use of Scripting Languages in Web Browsers and Email Clients	What percentage of the organization's hardware assets are utilizing unauthorized scripting languages that run in all web browsers and email clients?	
Maintain and Enforce Network-Based URL Filters	What percentage of the organization's hardware assets (whether physically at an organization's facilities or not) are not required to utilize network-based URL filters?	
Subscribe to URL-Categorization service	Has the organization subscribed to URL categorization services to ensure that they are up-to-date with the most recent website category definitions available?	
Log all URL requests	What percentage of the organization's hardware assets (whether physically at an organization's facilities or not) are not required to log all URL requests made from the organization's system?	
Use of DNS Filtering Services	What percentage of the organization's DNS servers are using DNS filtering to help block access to known malicious domains?	
Implement DMARC and Enable Receiver-Side Verification	Has the organization implemented Domain-based Message Authentication, Reporting and Conformance (DMARC), starting by implementing the Sender Policy Framework (SPF) and the DomainKeys Identified Mail(DKIM) standards?	
Block Unnecessary File Types	Has the organization blocked all e-mail attachments entering the organization's e-mail gateway if the file types are unnecessary for the organization's business?	
Sandbox All Email Attachments	Does the organization utilize sandboxing to analyze and block inbound email attachments with malicious behavior?	
Utilize Centrally Managed Anti-malware Software	What percentage of the organization's hardware assets do not utilize centrally managed anti-malware software to continuously monitor and defend each of the organization's workstations and servers?	
Ensure Anti-Malware Software and Signatures are Updated	What percentage of the organization's hardware assets do not utilize recently updated, centrally managed anti-malware software to continuously monitor and defend each of the organization's workstations and servers?	
Enable Operating System Anti-Exploitation Features/ Deploy Anti-Exploit Technologies	What percentage of the organization's hardware assets are not configured to require anti-exploitation features such as Data Execution Prevention (DEP) or Address Space Layout Randomization (ASLR) that are available in an operating system or deploy appropriate toolkits that can be configured to apply protection to a broader set of applications and executables?	
Configure Anti-Malware Scanning of Removable Devices	What percentage of the organization's hardware assets are not configured so that they automatically conduct an anti-malware scan of removable media when inserted or connected?	
Configure Devices Not To Auto-run Content	What percentage of the organization's hardware assets are not configured to not auto-run content from removable media?	
Centralize Anti-malware Logging	What percentage of the organization's hardware assets do not utilize centrally managed anti-malware software to continuously monitor and defend each of the organization's workstations and servers?	
Enable DNS Query Logging	What percentage of the organization's Domain Name System (DNS) servers are not configured to require query logging to detect hostname lookups for known malicious domains?	
Enable Command-line Audit Logging	What percentage of the organization's hardware assets have not enabled command-line audit logging for command shells, such as Python or Windows PowerShell with enhanced logging enabled?	
Associate Active Ports, Services and Protocols to Asset Inventory	What percentage of the organization's hardware assets do not associate active ports, services and protocols to the hardware assets in the asset inventory?	
Ensure Only Approved Ports, Protocols and Services Are Running	What percentage of the organization's hardware assets are not configured to require that only network ports, protocols, and services listening on a system with validated business needs, are running on each system?	
Perform Regular Automated Port Scans	What percentage of the organization's hardware assets are not regularly scanned by a port scanner to alert if unauthorized ports are detected on a system?	
Apply Host-based Firewalls or Port Filtering	What percentage of the organization's hardware assets are not utilizing host-based firewalls or port filtering tools on end systems, with a default-deny rule that drops all traffic except those services and ports that are explicitly allowed?	
Implement Application Firewalls	What percentage of the organization's critical servers are not required to utilize application layer firewalls to verify and validate the traffic going to the server?	
Ensure Regular Automated Back Ups	What percentage of the organization's hardware assets are not configured to back up system data automatically on a regular basis?	
Perform Complete System Backups	What percentage of the organization's hardware assets are not configured to back up the complete asset automatically on a regular basis?	
Test Data on Backup Media	What percentage of the organization's hardware asset backups have not been tested recently to ensure that the backup is working properly?	
Ensure Protection of Backups	What percentage of the organization's hardware asset backups are not properly protected via physical security or encryption when they are stored, as well as when they are moved across the network (this includes remote backups and cloud services as well)?	

Ensure Backups Have At least One Non-Continuously Addressable Destination	What percentage of the organization's hardware assets does not have at least one backup destination that is not continuously addressable through operating system calls?	
Maintain Standard Security Configurations for Network Devices	What percentage of the organization's network devices do not utilize a standard, documented security configuration standard for the device?	
Document Traffic Configuration Rules	What percentage of the organization's network devices do not have all configuration rules that allow traffic to flow through network devices be documented in a configuration management system with a specific business reason for each rule, a specific individual's name responsible for that business need, and an expected duration of the need?	
Use Automated Tools to Verify Standard Device Configurations and Detect Changes	What percentage of the organization's network devices are not regularly compared against approved security configurations defined for each network device in use and alert when any deviations are discovered?	
Install the Latest Stable Version of Any Security-related Updates on All Network Devices	What percentage of the organization's network devices are not utilizing the latest stable version of any security-related updates?	
Manage Network Devices Using Multi-Factor Authentication and Encrypted Sessions	What percentage of the organization's network devices are not managed using multi-factor authentication and encrypted sessions?	
Use Dedicated Machines For All Network Administrative Tasks	What percentage of the organization's network engineers are not utilizing a dedicated machine for all administrative tasks or tasks requiring elevated access to the organization's network devices?	
Manage Network Infrastructure Through a Dedicated Network	What percentage of the organization's network engineers are not utilizing a dedicated machine, located on a dedicated management network, for all administrative tasks or tasks requiring elevated access to the organization's network devices?	
Maintain an Inventory of Network Boundaries	Does the organization maintain an up-to-date inventory of all of the organization's network boundaries?	
Scan for Unauthorized Connections across Trusted Network Boundaries	What percentage of the organization's hardware assets have not recently been scanned to identify unauthorized network boundaries?	
Deny Communications with Known Malicious IP Addresses	Are each of the organization's network boundaries configured to deny communications with known malicious or unused Internet IP addresses and limit access only to trusted and necessary IP address ranges?	
Deny Communication over Unauthorized Ports	Are each of the organization's network boundaries configured to deny communication over unauthorized TCP or UDP ports or application traffic to ensure that only authorized protocols are allowed to cross the network boundary in or out of the network?	
Configure Monitoring Systems to Record Network Packets	What percentage of the organization's network boundaries are not configured to record network packets passing through the boundary?	
Deploy Network-based IDS Sensor	What percentage of the organization's network boundaries are not configured to require network-based Intrusion Detection Systems (IDS) sensors to look for unusual attack mechanisms and detect compromise of these systems the boundary?	
Deploy Network-Based Intrusion Prevention Systems	What percentage of the organization's organization's network boundaries are not configured to require network-based Intrusion Prevention Systems (IPS) sensors to look for unusual attack mechanisms and detect compromise of these systems the boundary?	
Deploy NetFlow Collection on Networking Boundary Devices	What percentage of the organization's network boundary devices are not required to use NetFlow and logging data on the devices?	
Deploy Application Layer Filtering Proxy Server	What percentage of the organization's network boundaries are not configured to pass through an authenticated application layer proxy that is configured to filter unauthorized connections?	
Decrypt Network Traffic at Proxy	What percentage of the organization's network boundaries are not configured to decrypt all encrypted network traffic prior to analyzing the content?	
Require All Remote Login to Use Multi-factor Authentication	What percentage of the organization's hardware devices are not required to utilize encryption and multi-factor authentication when remotely accessing the organization's network systems?	
Manage All Devices Remotely Logging into Internal Network	What percentage of the organization's devices remotely logging into the organization's network are not scanned prior to accessing the network to ensure that each of the organization's security policies has been enforced in the same manner as local network devices?	
Maintain an Inventory Sensitive Information	Does the organization maintain an inventory of all sensitive information stored, processed, or transmitted by the organization's technology systems, including those located onsite or at a remote service provider?	
Remove Sensitive Data or Systems Not Regularly Accessed by Organization	Does the organization regularly remove sensitive data sets or systems not regularly accessed by the organization from the network?	
Monitor and Block Unauthorized Network Traffic	Has the organization deployed an automated tool on network perimeters that monitors for sensitive information and blocks such transfers while alerting information security professionals?	
Only Allow Access to Authorized Cloud Storage or Email Providers	Does the organization only allow access to authorized cloud storage or email providers?	

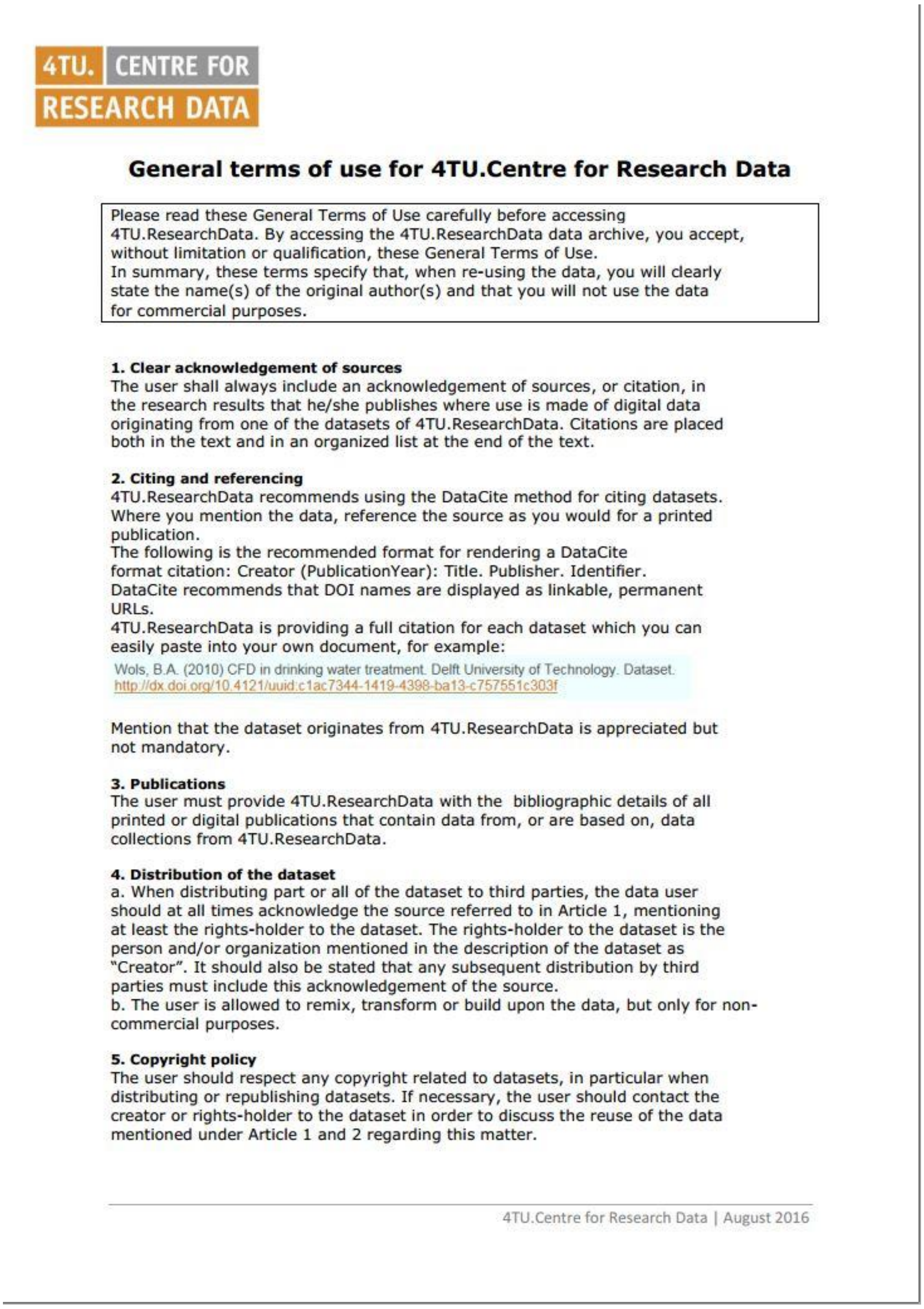
Monitor and Detect Any Unauthorized Use of Encryption	What percentage of the organization's network boundaries are not configured to monitor all traffic leaving the organization and detect any unauthorized use of encryption?	
Encrypt the Hard Drive of All Mobile Devices.	What percentage of the organization's mobile devices do not utilize approved whole disk encryption software?	
Manage USB Devices	What percentage of the organization's hardware assets are not configured to only allow the use of specific USB devices?	
Manage System's External Removable Media's Read/write Configurations	What percentage of the organization's hardware assets are not configured not to write data to USB storage devices, if there is no business need for supporting such devices?	
Encrypt Data on USB Storage Devices	What percentage of the organization's hardware assets are not configured to encrypt all data stored on USB devices?	
Segment the Network Based on Sensitivity	What percentage of the organization's network devices are not located on dedicated Virtual Local Area Networks (VLANs)?	
Enable Firewall Filtering Between VLANs	What percentage of the organization's network devices are not located on dedicated Virtual Local Area Networks (VLANs) separated by firewall filters?	
Disable Workstation to Workstation Communication	What percentage of the organization's workstation devices are not located on dedicated Private Virtual Local Area Networks (PVLANs)?	
Encrypt All Sensitive Information in Transit	What percentage of the organization's sensitive information is not encrypted in transit?	
Utilize an Active Discovery Tool to Identify Sensitive Data	What percentage of the organization's assets have not been scanned by an active discovery tool to identify all sensitive information stored, processed, or transmitted by the organization's technology systems?	
Protect Information through Access Control Lists	What percentage of the organization's hardware assets have not been configured with appropriate file system, network share, claims, application, or database specific access control lists?	
Enforce Access Control to Data through Automated Tools	What percentage of the organizations systems do not use an automated tool, such as host-based Data Loss Prevention, to enforce access controls to data even when data is copied off a system?	
Encrypt Sensitive Information at Rest	What percentage of the organization's sensitive information is not encrypted at rest and requires a secondary authentication mechanism not integrated into the operating system, in order to access the information?	
Enforce Detail Logging for Access or Changes to Sensitive Data	What percentage of the organization's sensitive information does not require detailed audit logging when the data is accessed?	
Maintain an Inventory of Authorized Wireless Access Points	What percentage of the organization's wireless access points have not been authorized in the organization's wireless access point inventory?	
Detect Wireless Access Points Connected to the Wired Network	What percentage of the organization's hardware assets have not recently been scanned to detect and alert on unauthorized wireless access points connected to the wired network?	
Use a Wireless Intrusion Detection System	What percentage of the organization's facilities do not have a wireless intrusion detection system (WIDS) to detect and alert on unauthorized wireless access points connected to the network?	
Disable Wireless Access on Devices if Not Required	What percentage of the organization's hardware assets is not configured to disable wireless access in devices that do not have a business purpose for wireless access?	
Limit Wireless Access on Client Devices	What percentage of the organization's hardware assets are not configured to allow access only to authorized wireless networks and to restrict access for other wireless networks?	
Disable Peer-to-peer Wireless Network Capabilities on Wireless Clients	What percentage of the organization's hardware assets are not configured to disable peer-to-peer (adhoc) wireless network capabilities on wireless clients?	
Leverage the Advanced Encryption Standard (AES) to Encrypt Wireless Data	What percentage of the organization's hardware assets are not configured to leverage the Advanced Encryption Standard (AES) to encrypt wireless data in transit?	
Use Wireless Authentication Protocols that Require Mutual, Multi-Factor Authentication	What percentage of the organization's hardware assets are not configured to utilize wireless networks to use authentication protocols such as Extensible Authentication Protocol-Transport Layer Security (EAP/TLS), that requires mutual, multi-factor authentication?	
Disable Wireless Peripheral Access of Devices	What percentage of the organization's hardware assets are not configured to disable wireless peripheral access of devices (such as Bluetooth), unless such access is required for a business purpose?	
Create Separate Wireless Network for Personal and Untrusted Devices	Does the organization utilize a separate a wireless network for personal or untrusted devices?	
Maintain an Inventory of Authentication Systems	What percentage of the organization's authentication systems are not included in the organization's inventory?	

Configure Centralized Point of Authentication	Has the organization configured access for all accounts through as few centralized points of authentication as possible, including network, security, and cloud systems?	
Require Multi-factor Authentication	What percentage of the organization's user accounts do not require multi-factor authentication?	
Encrypt or Hash all Authentication Credentials	What percentage of the organization's hardware assets' authentication files cannot be accessed without root or administrator privileges and are not encrypted or hashed?	
Encrypt Transmittal of Username and Authentication Credentials	What percentage of the organization's user accounts and authentication credentials are not transmitted across networks using encrypted channels?	
Maintain an Inventory of Accounts	What percentage of the organization's accounts are not included in the organization's inventory?	
Establish Process for Revoking Access	Has the organization established and followed an automated process for revoking system access by disabling accounts immediately upon termination or change of responsibilities of an employee or contractor?	
Disable Any Unassociated Accounts	What percentage of the organization's user accounts are not disabled if they cannot be associated with a business process or owner?	
Disable Dormant Accounts	Does the organization automatically disable dormant accounts after a set period of inactivity?	
Ensure All Accounts Have An Expiration Date	What percentage of the organization's user accounts do not have an expiration date that is monitored and enforced?	
Lock Workstation Sessions After Inactivity	Does the organization automatically lock workstation sessions after a standard period of inactivity?	
Monitor Attempts to Access Deactivated Accounts	Does the organization monitor attempts to access deactivated accounts through audit logging?	
Alert on Account Login Behavior Deviation	Does the organization alert when users deviate from normal login behavior, such as time-of-day, workstation location and duration?	
Perform a Skills Gap Analysis	Has the organization performed a skills gap analysis to understand the skills and behaviors workforce members are not adhering to, using this information to build a baseline education roadmap.	
Deliver Training to Fill the Skills Gap	Has the organization delivered training to address the skills gap identified to positively impact workforce members' security behavior.	
Implement a Security Awareness Program	Has the organization created a security awareness program for all workforce members to complete on a regular basis to ensure they understand and exhibit the necessary behaviors and skills to help ensure the security of the organization. The organization's security awareness program should be communicated in a continuous and engaging manner.	
Update Awareness Content Frequently	Has the organization ensured that the organization's security awareness program is updated frequently (at least annually) to address new technologies, threats, standards and business requirements.	
Train Workforce on Secure Authentication	Has the organization trained workforce members on the importance of enabling and utilizing secure authentication.	
Train Workforce on Identifying Social Engineering Attacks	Has the organization trained the workforce on how to identify different forms of social engineering attacks, such as phishing, phone scams and impersonation calls.	
Train Workforce on Sensitive Data Handling	Has the organization trained workforce on how to identify and properly store, transfer, archive and destroy sensitive information.	
Train Workforce on Causes of Unintentional Data Exposure	Has the organization trained workforce members to be aware of causes for unintentional data exposures, such as losing their mobile devices or emailing the wrong person due to autocomplete in email.	
Train Workforce Members on Identifying and Reporting Incidents	Has the organization trained employees to be able to identify the most common indicators of an incident and be able to report such an incident.	
Establish Secure Coding Practices	Has the organization established secure coding practices appropriate to the programming language and development environment being used.	
Ensure Explicit Error Checking is Performed for All In-house Developed Software	For in-house developed software, has the organization ensured that explicit error checking is performed and documented for all input, including for size, data type, and acceptable ranges or formats.	
Verify That Acquired Software is Still Supported	Has the organization verified that the version of all software acquired from outside your organization is still supported by the developer or appropriately hardened based on developer security recommendations.	
Only Use Up-to-date And Trusted Third-Party Components	Has the organization only used up-to-date and trusted third-party components for the software developed by the organization.	
Use Only Standardized and Extensively Reviewed Encryption Algorithms	Has the organization used only standardized and extensively reviewed encryption algorithms.	

Ensure Software Development Personnel are Trained in Secure Coding	Has the organization ensured that all software development personnel receive training in writing secure code for their specific development environment and responsibilities.	
Apply Static and Dynamic Code Analysis Tools	Has the organization applied static and dynamic analysis tools to verify that secure coding practices are being adhered to for internally developed software.	
Establish a Process to Accept and Address Reports of Software Vulnerabilities	Has the organization established a process to accept and address reports of software vulnerabilities, including providing a means for external entities to contact your security group.	
Separate Production and Non-Production Systems	Has the organization maintained separate environments for production and nonproduction systems. Developers should not have unmonitored access to production environments.	
Deploy Web Application Firewalls (WAFs)	Has the organization protected web applications by deploying web application firewalls (WAFs) that inspect all traffic flowing to the web application for common web application attacks. For applications that are not web-based, specific application firewalls should be deployed if such tools are available for the given application type. If the traffic is encrypted, the device should either sit behind the encryption or be capable of decrypting the traffic prior to analysis. If neither option is appropriate, a host-based web application firewall should be deployed.	
Use Standard Hardening Configuration Templates for Databases	For applications that rely on a database, has the organization used standard hardening configuration templates. All systems that are part of critical business processes should also be tested.	
Document Incident Response Procedures	Has the organization ensured that there are written incident response plans that defines roles of personnel as well as phases of incident handling/management.	
Assign Job Titles and Duties for Incident Response	Has the organization assigned job titles and duties for handling computer and network incidents to specific individuals and ensure tracking and documentation throughout the incident through resolution.	
Designate Management Personnel to Support Incident Handling	Has the organization designated management personnel, as well as backups, who will support the incident handling process by acting in key decision-making roles.	
Devise Organization-wide Standards for Reporting Incidents	Has the organization devised organization-wide standards for the time required for system administrators and other workforce members to report anomalous events to the incident handling team, the mechanisms for such reporting, and the kind of information that should be included in the incident notification.	
Maintain Contact Information For Reporting Security Incidents	Has the organization assembled and maintain information on third-party contact information to be used to report a security incident, such as Law Enforcement, relevant government departments, vendors, and ISAC partners.	
Publish Information Regarding Reporting Computer Anomalies and Incidents	Has the organization published information for all workforce members, regarding reporting computer anomalies and incidents to the incident handling team. Such information should be included in routine employee awareness activities.	
Conduct Periodic Incident Scenario Sessions for Personnel	Has the organization planned and conducted routine incident response exercises and scenarios for the workforce involved in the incident response to maintain awareness and comfort in responding to real world threats. Exercises should test communication channels, decision making, and incident responders technical capabilities using tools and data available to them.	
Create Incident Scoring and Prioritization Schema	Has the organization created incident scoring and prioritization schema based on known or potential impact to your organization. Utilize score to define frequency of status updates and escalation procedures.	
Establish a Penetration Testing Program	Has the organization established a program for penetration tests that includes a full scope of blended attacks, such as wireless, client-based, and web application attacks.	
Conduct Regular External and Internal Penetration Tests	Has the organization conducted regular external and internal penetration tests to identify vulnerabilities and attack vectors that can be used to exploit enterprise systems successfully.	
Perform Periodic Red Team Exercises	Has the organization performed periodic Red Team exercises to test organizational readiness to identify and stop attacks or to respond quickly and effectively.	
Include Tests for Presence of Unprotected System Information and Artifacts	Has the organization included tests for the presence of unprotected system information and artifacts that would be useful to attackers, including network diagrams, configuration files, older penetration test reports, e-mails or documents containing passwords or other information critical to system operation.	
Create Test Bed for Elements Not Typically Tested in Production	Has the organization created a test bed that mimics a production environment for specific penetration tests and Red Team attacks against elements that are not typically tested in production, such as attacks against supervisory control and data acquisition and other control systems.	
Use Vulnerability Scanning and Penetration Testing Tools in Concert	Has the organization used vulnerability scanning and penetration testing tools in concert. The results of vulnerability scanning assessments should be used as a starting point to guide and focus penetration testing efforts.	
Ensure Results from Penetration Test are Documented Using Open, Machine-readable Standards	Has the organization, wherever possible, ensured that Red Teams results are documented using open, machine-readable standards (e.g., SCAP). Devise a scoring method for determining the results of Red Team exercises so that results can be compared over time.	
Control and Monitor Accounts Associated with Penetration Testing	Has the organization ensured that any user or system accounts used to perform penetration testing should be controlled and monitored to make sure they are only being used for legitimate purposes, and are removed or restored to normal function after testing is over.	

9.
Appendix B – Data Set GTU

General Terms of Use of 4TU Data Set:



4TU. CENTRE FOR RESEARCH DATA

General terms of use for 4TU.Centre for Research Data

Please read these General Terms of Use carefully before accessing 4TU.ResearchData. By accessing the 4TU.ResearchData data archive, you accept, without limitation or qualification, these General Terms of Use.
In summary, these terms specify that, when re-using the data, you will clearly state the name(s) of the original author(s) and that you will not use the data for commercial purposes.

1. Clear acknowledgement of sources
The user shall always include an acknowledgement of sources, or citation, in the research results that he/she publishes where use is made of digital data originating from one of the datasets of 4TU.ResearchData. Citations are placed both in the text and in an organized list at the end of the text.

2. Citing and referencing
4TU.ResearchData recommends using the DataCite method for citing datasets. Where you mention the data, reference the source as you would for a printed publication.
The following is the recommended format for rendering a DataCite format citation: Creator (PublicationYear): Title. Publisher. Identifier.
DataCite recommends that DOI names are displayed as linkable, permanent URLs.
4TU.ResearchData is providing a full citation for each dataset which you can easily paste into your own document, for example:

Wols, B.A. (2010) CFD in drinking water treatment. Delft University of Technology. Dataset.
http://dx.doi.org/10.4121/uuid:c1ac7344-1419-4398-ba13-c757551c303f

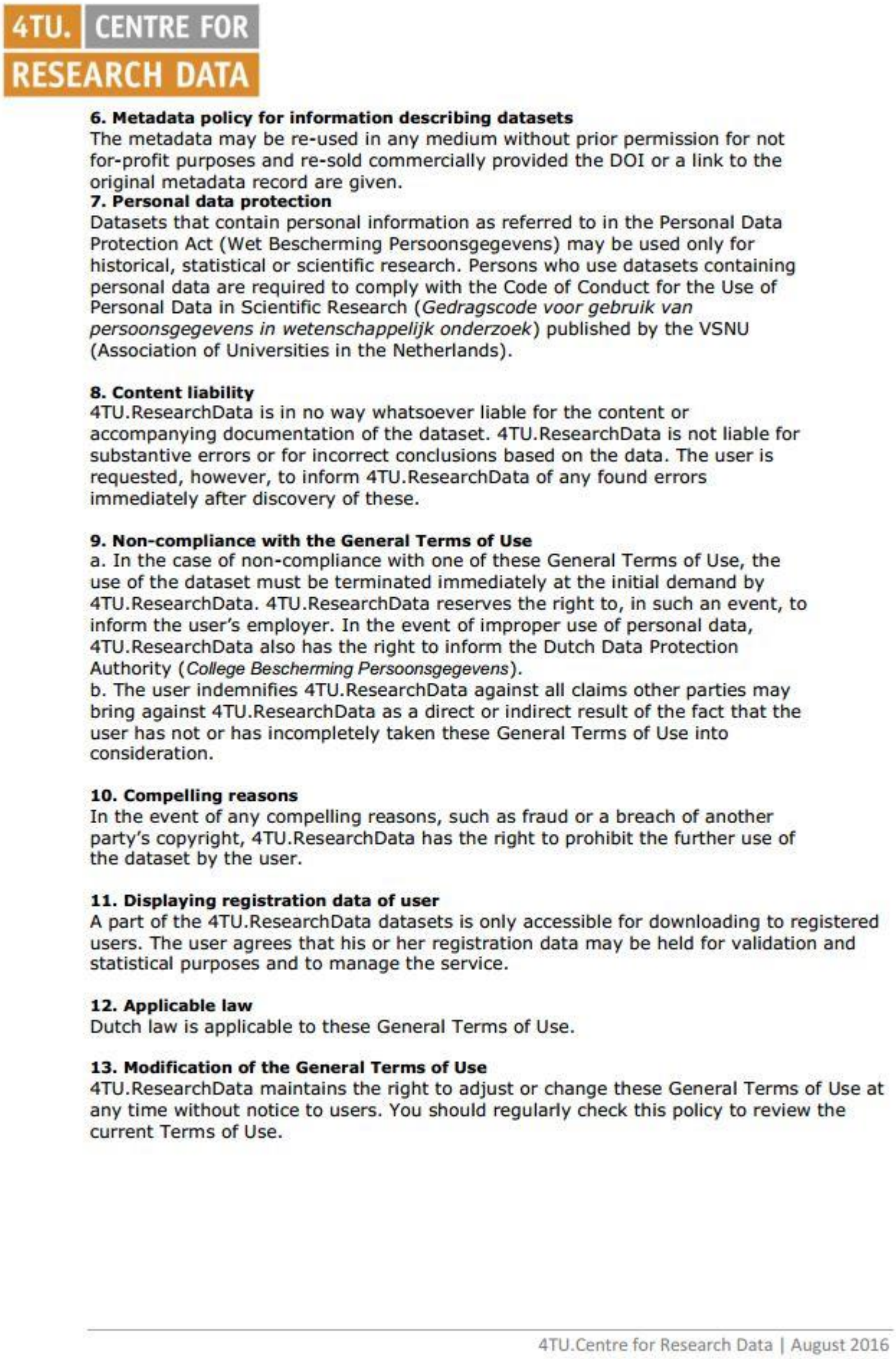
Mention that the dataset originates from 4TU.ResearchData is appreciated but not mandatory.

3. Publications
The user must provide 4TU.ResearchData with the bibliographic details of all printed or digital publications that contain data from, or are based on, data collections from 4TU.ResearchData.

4. Distribution of the dataset
a. When distributing part or all of the dataset to third parties, the data user should at all times acknowledge the source referred to in Article 1, mentioning at least the rights-holder to the dataset. The rights-holder to the dataset is the person and/or organization mentioned in the description of the dataset as "Creator". It should also be stated that any subsequent distribution by third parties must include this acknowledgement of the source.
b. The user is allowed to remix, transform or build upon the data, but only for non-commercial purposes.

5. Copyright policy
The user should respect any copyright related to datasets, in particular when distributing or republishing datasets. If necessary, the user should contact the creator or rights-holder to the dataset in order to discuss the reuse of the data mentioned under Article 1 and 2 regarding this matter.

4TU.Centre for Research Data | August 2016



4TU. CENTRE FOR RESEARCH DATA

6. Metadata policy for information describing datasets
The metadata may be re-used in any medium without prior permission for not for-profit purposes and re-sold commercially provided the DOI or a link to the original metadata record are given.

7. Personal data protection
Datasets that contain personal information as referred to in the Personal Data Protection Act (Wet Bescherming Persoonsgegevens) may be used only for historical, statistical or scientific research. Persons who use datasets containing personal data are required to comply with the Code of Conduct for the Use of Personal Data in Scientific Research (*Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek*) published by the VSNU (Association of Universities in the Netherlands).

8. Content liability
4TU.ResearchData is in no way whatsoever liable for the content or accompanying documentation of the dataset. 4TU.ResearchData is not liable for substantive errors or for incorrect conclusions based on the data. The user is requested, however, to inform 4TU.ResearchData of any found errors immediately after discovery of these.

9. Non-compliance with the General Terms of Use
a. In the case of non-compliance with one of these General Terms of Use, the use of the dataset must be terminated immediately at the initial demand by 4TU.ResearchData. 4TU.ResearchData reserves the right to, in such an event, to inform the user's employer. In the event of improper use of personal data, 4TU.ResearchData also has the right to inform the Dutch Data Protection Authority (*College Bescherming Persoonsgegevens*).
b. The user indemnifies 4TU.ResearchData against all claims other parties may bring against 4TU.ResearchData as a direct or indirect result of the fact that the user has not or has incompletely taken these General Terms of Use into consideration.

10. Compelling reasons
In the event of any compelling reasons, such as fraud or a breach of another party's copyright, 4TU.ResearchData has the right to prohibit the further use of the dataset by the user.

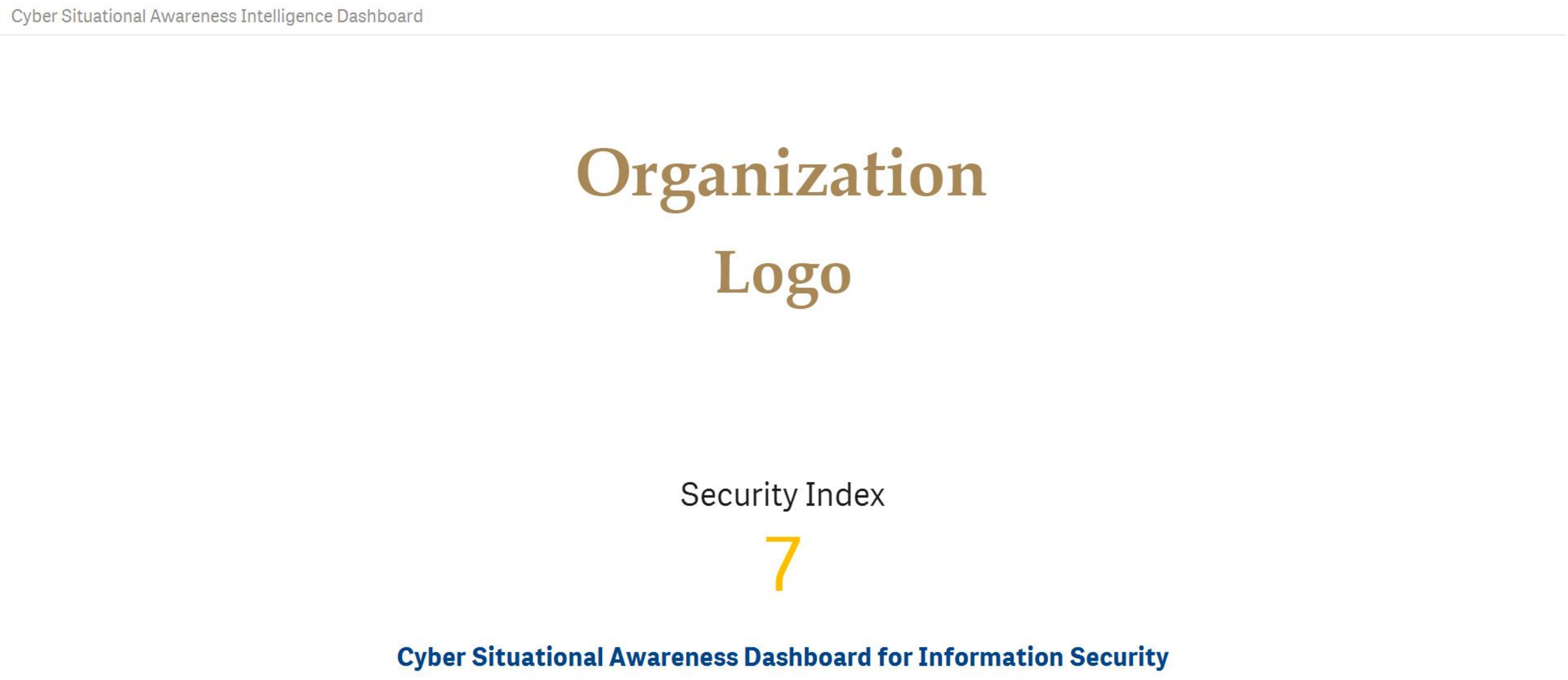
11. Displaying registration data of user
A part of the 4TU.ResearchData datasets is only accessible for downloading to registered users. The user agrees that his or her registration data may be held for validation and statistical purposes and to manage the service.

12. Applicable law
Dutch law is applicable to these General Terms of Use.

13. Modification of the General Terms of Use
4TU.ResearchData maintains the right to adjust or change these General Terms of Use at any time without notice to users. You should regularly check this policy to review the current Terms of Use.

4TU.Centre for Research Data | August 2016

10.
Appendix C – Dashboard Views



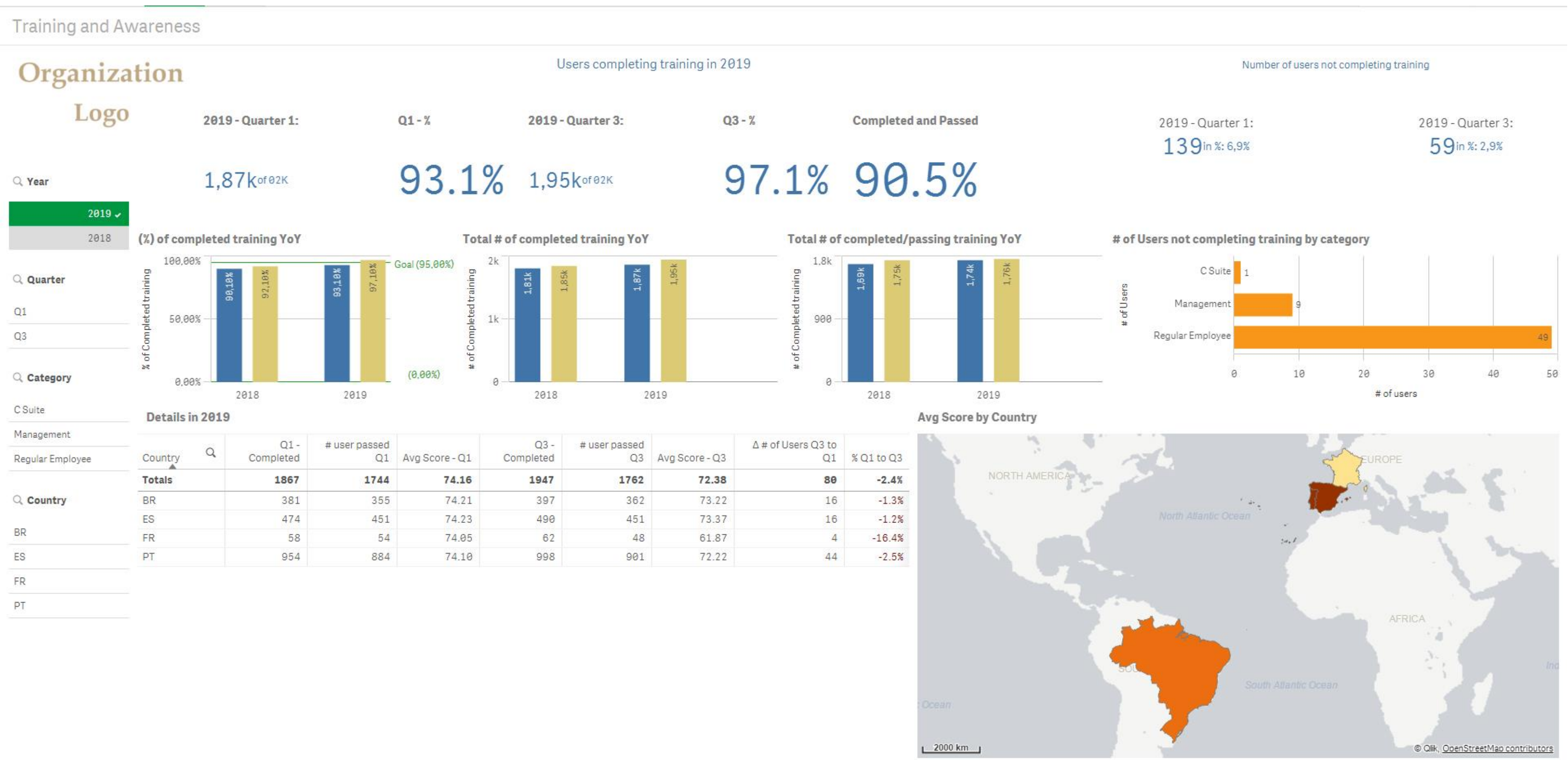
Cyber Situational Awareness Intelligence Dashboard

Organization
Logo

Security Index

7

Cyber Situational Awareness Dashboard for Information Security



Training and Awareness

Organization Logo

Users completing training in 2019

2019 - Quarter 1: 1,87k of 02K

Q1 - % 93.1%

2019 - Quarter 3: 1,95k of 02K

Q3 - % 97.1%

Completed and Passed 90.5%

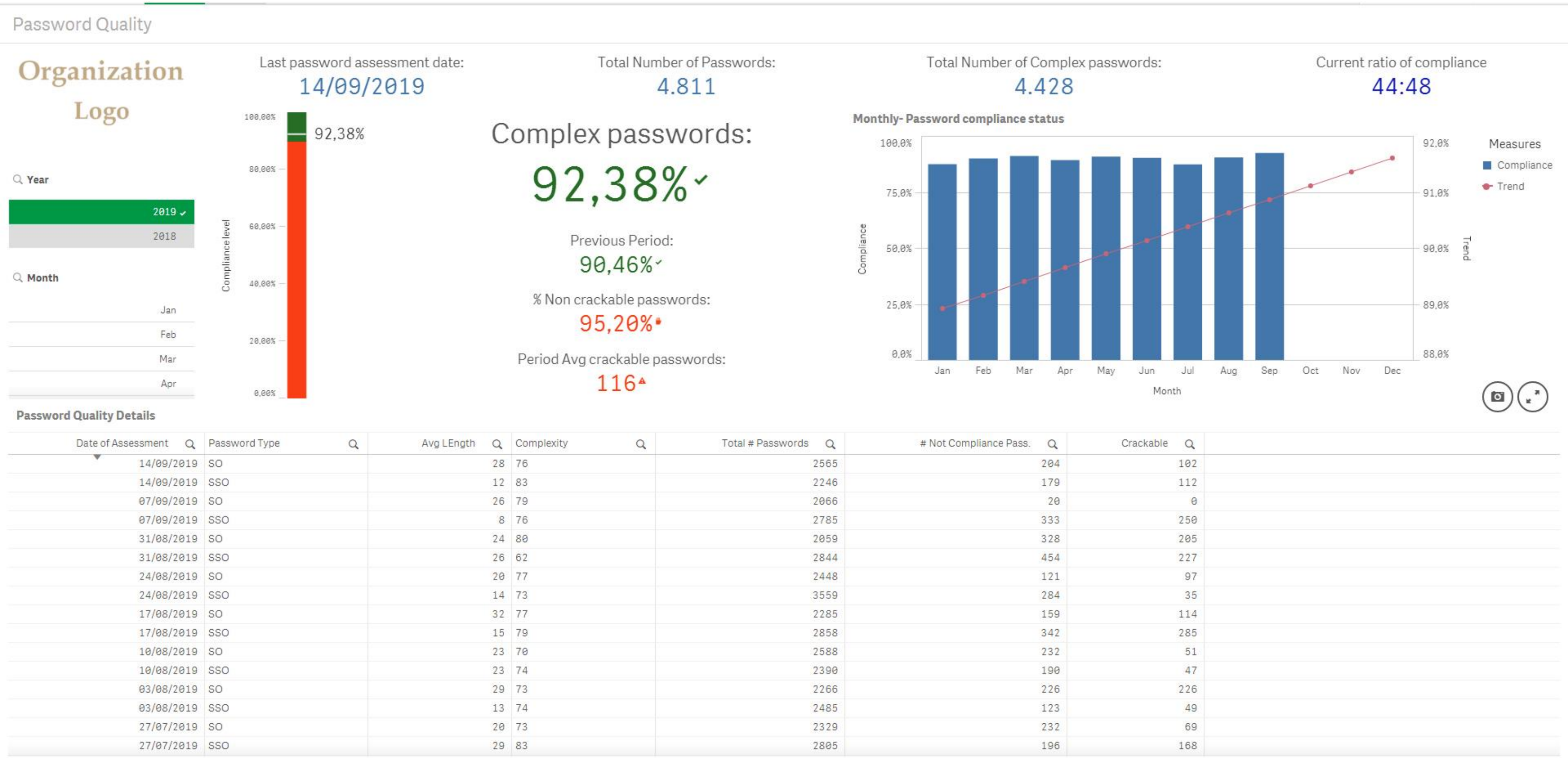
Number of users not completing training

2019 - Quarter 1: 139 in %: 6,9%

2019 - Quarter 3: 59 in %: 2,9%

Details in 2019

Country	Q1 - Completed	# user passed Q1	Avg Score - Q1	Q3 - Completed	# user passed Q3	Avg Score - Q3	Δ # of Users Q3 to Q1	% Q1 to Q3
Totals	**1867**	**1744**	**74.16**	**1947**	**1762**	**72.38**	**80**	**-2.4%**
BR	381	355	74.21	397	362	73.22	16	-1.3%
ES	474	451	74.23	490	451	73.37	16	-1.2%
FR	58	54	74.05	62	48	61.87	4	-16.4%
PT	954	884	74.10	998	901	72.22	44	-2.5%



Last password assessment date:
14/09/2019

Total Number of Passwords:
4.811

Total Number of Complex passwords:
4.428

Current ratio of compliance
44:48

Complex passwords:
92,38% ✓

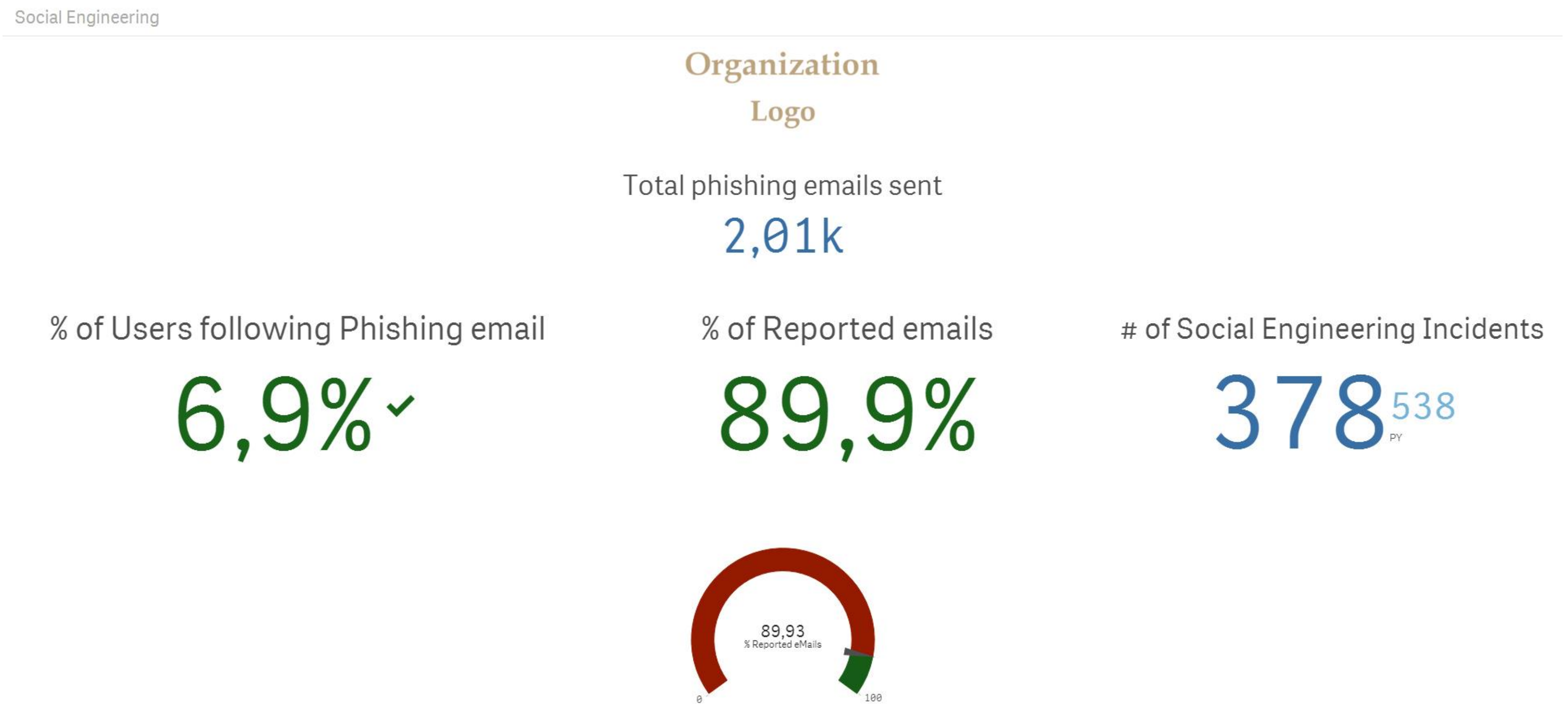
Previous Period:
90,46% ✓

% Non crackable passwords:
95,20% ●

Period Avg crackable passwords:
116 ▲

Password Quality Details

Date of Assessment	Password Type	Avg LEngth	Complexity	Total # Passwords	# Not Compliance Pass.	Crackable
14/09/2019	SO	28	76	2565	204	102
14/09/2019	SSO	12	83	2246	179	112
07/09/2019	SO	26	79	2066	20	0
07/09/2019	SSO	8	76	2785	333	250
31/08/2019	SO	24	80	2059	328	205
31/08/2019	SSO	26	62	2844	454	227
24/08/2019	SO	20	77	2448	121	97
24/08/2019	SSO	14	73	3559	284	35
17/08/2019	SO	32	77	2285	159	114
17/08/2019	SSO	15	79	2858	342	285
10/08/2019	SO	23	70	2588	232	51
10/08/2019	SSO	23	74	2390	190	47
03/08/2019	SO	29	73	2266	226	226
03/08/2019	SSO	13	74	2485	123	49
27/07/2019	SO	20	73	2329	232	69
27/07/2019	SSO	29	83	2805	196	168



Social Engineering

Organization
Logo

Total phishing emails sent

2,01k

% of Users following Phishing email

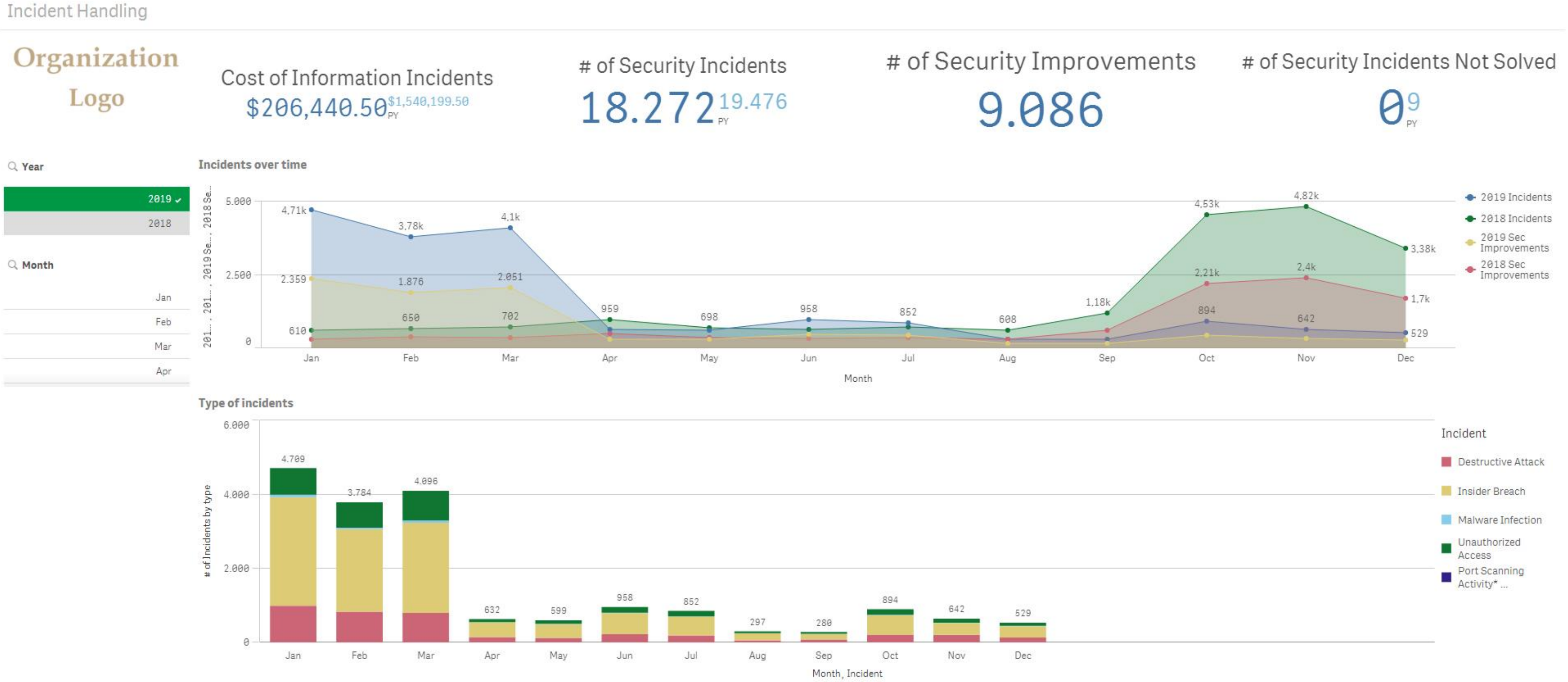
6,9% ✓

% of Reported emails

89,9%

# of Social Engineering Incidents

378 538 PY



Incident Handling

Organization Logo

Cost of Information Incidents

$206,440.50 $1,540,199.50 PY

# of Security Incidents

18.272 19.476 PY

# of Security Improvements

9.086

# of Security Incidents Not Solved

0 9 PY

Year

2019 ✓

2018

Month

Jan

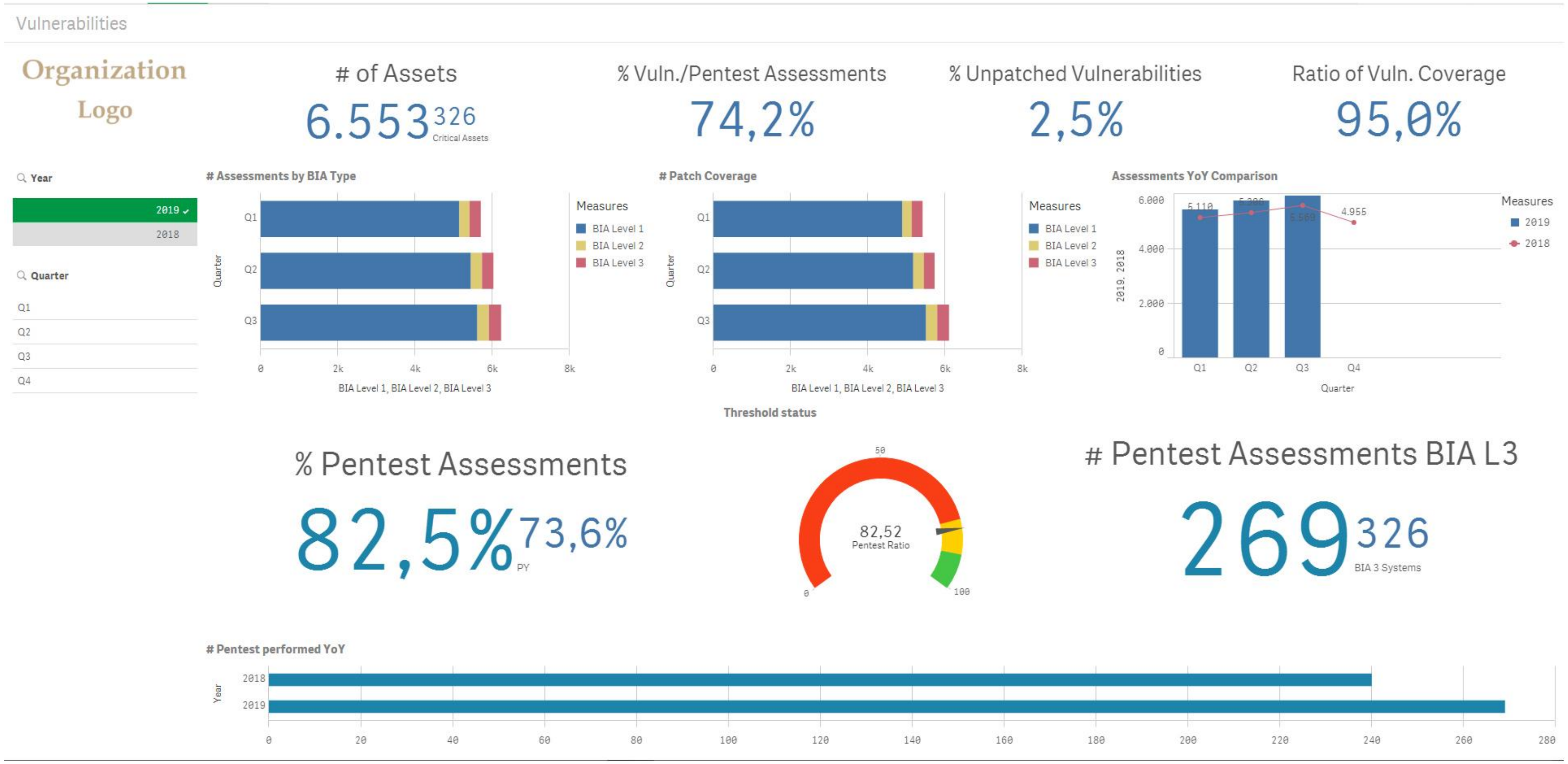
Feb

Mar

Apr

Incidents over time

Type of incidents



Vulnerabilities

Organization Logo

Year: 2019 ✓ | 2018

Quarter: Q1 | Q2 | Q3 | Q4

# of Assets: 6.553 — 326 Critical Assets

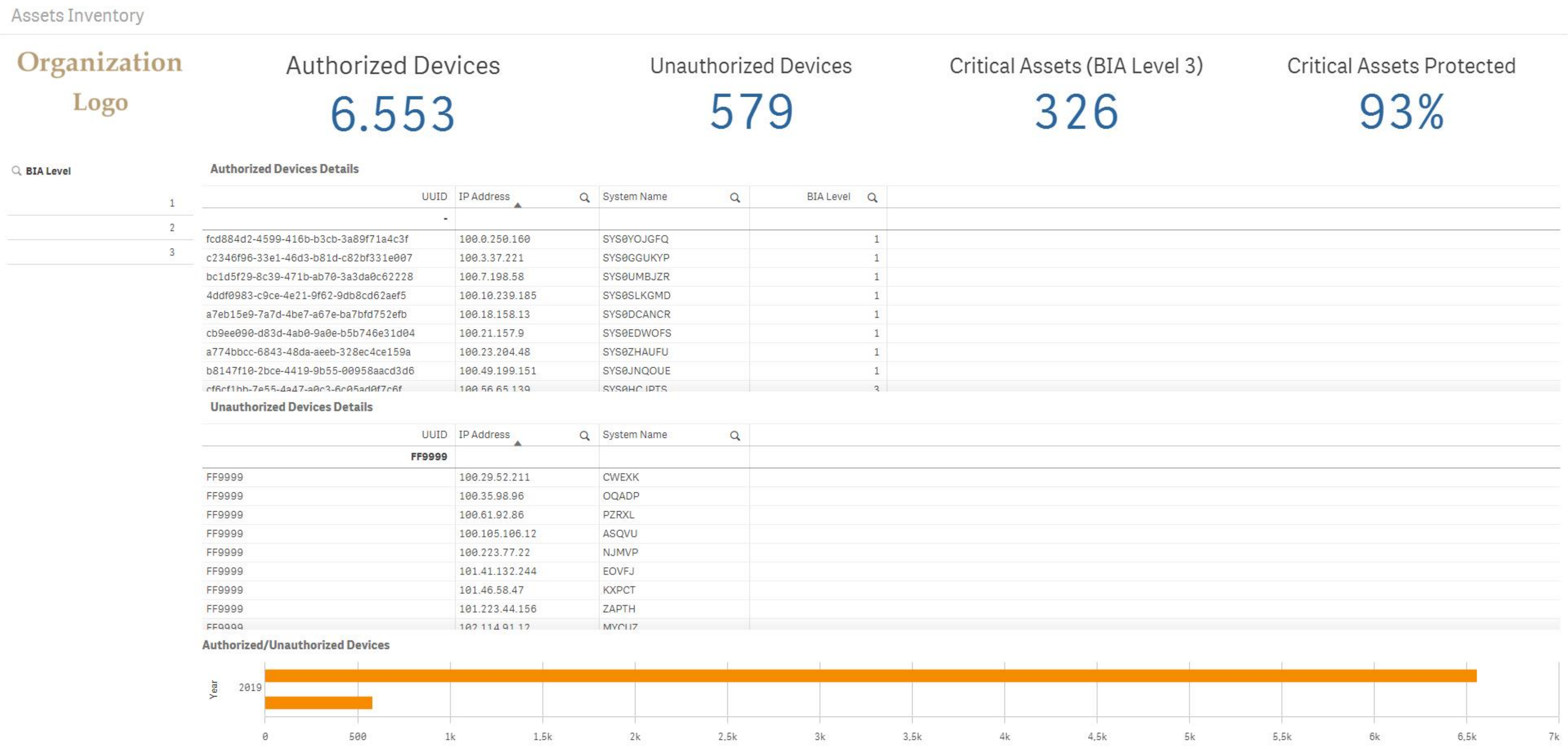
% Vuln./Pentest Assessments: 74,2%

% Unpatched Vulnerabilities: 2,5%

Ratio of Vuln. Coverage: 95,0%

% Pentest Assessments: 82,5% — 73,6% PY

Pentest Assessments BIA L3: 269 — 326 BIA 3 Systems



Assets Inventory

Organization Logo

Authorized Devices	Unauthorized Devices	Critical Assets (BIA Level 3)	Critical Assets Protected
6.553	579	326	93%

BIA Level

1

2

3

Authorized Devices Details

UUID	IP Address	System Name	BIA Level
-			
fcd884d2-4599-416b-b3cb-3a89f71a4c3f	100.0.250.160	SYS0YOJGFQ	1
c2346f96-33e1-46d3-b81d-c82bf331e007	100.3.37.221	SYS0GGUKYP	1
bc1d5f29-8c39-471b-ab70-3a3da0c62228	100.7.198.58	SYS0UMBJZR	1
4ddf0983-c9ce-4e21-9f62-9db8cd62aef5	100.10.239.185	SYS0SLKGMD	1
a7eb15e9-7a7d-4be7-a67e-ba7bfd752efb	100.18.158.13	SYS0DCANCR	1
cb9ee090-d83d-4ab0-9a0e-b5b746e31d04	100.21.157.9	SYS0EDWOFS	1
a774bbcc-6843-48da-aeeb-328ec4ce159a	100.23.204.48	SYS0ZHAUFU	1
b8147f10-2bce-4419-9b55-00958aacd3d6	100.49.199.151	SYS0JNQOUE	1
cf6cf1bb-7e55-4a47-a0c3-6c05ad0f7c6f	100.56.65.139	SYS0HCJPTS	3

Unauthorized Devices Details

UUID	IP Address	System Name
FF9999		
FF9999	100.29.52.211	CWEXK
FF9999	100.35.98.96	OQADP
FF9999	100.61.92.86	PZRXL
FF9999	100.105.106.12	ASQVU
FF9999	100.223.77.22	NJMVP
FF9999	101.41.132.244	EOVFJ
FF9999	101.46.58.47	KXPCT
FF9999	101.223.44.156	ZAPTH
FF9999	102.114.91.12	MYCUZ

Authorized/Unauthorized Devices

11.

Appendix D – Set Analysis

Metric Name	Set Analysis
# users Completing Training Q1	sum({<[_typeFact]={1}, Month=, Day=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"} >}Completed)
# users Q1	=' of '&chr(23)&num(count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)/1000,'##.00')&'K'
# users Completing Training Q3	sum({<[_typeFact]={1}, Month=, Day=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"} >}Completed)
# users Q3	=' of '&chr(23)&num(count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)/1000,'##.00')&'K'
% Users completing training Q1	(count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID) /count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID))
% Users completing training Q3	(count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID) /count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID))
% Users completing training YoY	num((sum({<[_typeFact]={1},Year=, Month=, Day=>}Completed)/count({<[_typeFact]={1},Year=, Month=, Day=, Completed={0,1} >}UserID)),'##,#0%')
# Users completing training YoY	sum({<Year=, [_typeFact]={1}, Month=, Day=>}Completed)
# Type of users not completing training	count({<[_typeFact]={1}, B25, Year=, Quarter=, Month=, Completed={0} >}distinct UserID)
# users not completing training Q1	=' in %: '&num(1-(count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)),'##,#%')

% users not completing training Q1	=' in %: '&num(1-(count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)),'##,#%')
# users not completing training Q3	=' in %: '&num(1-(count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)),'##,#%')
% users not completing training Q3	=' in %: '&num(1-(count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)/count({<[_typeFact]={1}, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={0,1} >}UserID)),'##,#%')
# users not completing training by Category	=count({<[_typeFact]={1}, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={0} >}UserID)
Users completing and passing training	(sum({<[_typeFact]={1}, PeriodID = {$(vMaxTAPeriod)},Year = {"$(=Year(Max(Date)))"}, Completed={1}>}_hasPassed) /sum({<[_typeFact]={1}, PeriodID = {$(vMaxTAPeriod)},Year = {"$(=Year(Max(Date)))"} >}Completed))
Table: # Completed Q1	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)
Table: # Completed Q3	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)
Table: # Passed Q1	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1}, [_hasPassed]={1} >}UserID)
Table: # Passed Q3	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1}, [_hasPassed]={1} >}UserID)
Table: AVG Score Q1	Avg({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}Score)
Table: AVG Score Q3	Avg({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}Score)
Table: Δ # of Users Q3 to Q1	count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)-count({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}UserID)
Table: Δ % Q1 to Q3	(Avg({<[_typeFact]={1}, Day=, Month=, Quarter={'Q3'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}Score)/Avg({<[_typeFact]={1}, Day=, Month=, Quarter={'Q1'},Year = {"$(=Year(Max(Date)))"}, Completed={1} >}Score)-1)

Map: Average Score by country	Avg({<[_typeFact]={1}, Day=, Month=, Quarter=,Year =, PeriodID = {$(vMaxTAPeriod)}, Completed={1} >}Score
Total phishing emails sent	count({<_typeFact={2}, Year = {"$(=Year(Max(Date)))"}, PeriodID = {$(vMaxPeriodSE)} >}UserID)
% of Users following Phishing email	=(count({<[_typeFact]={2}, Followed={1}, PeriodID = {$(vMaxPeriodSE)}>}UserID)/count({<[_typeFact]={2}, Followed={1,0}, PeriodID = {$(vMaxPeriodSE)}>}UserID))
% of Reported emails	(count({<[_typeFact]={2}, Reported={1}, PeriodID = {$(vMaxPeriodSE)}>}UserID)/count({<[_typeFact]={2}, Reported={1,0}, PeriodID = {$(vMaxPeriodSE)}>}UserID))
# of Social Engineering emails	num(sum({<Year = {"$(=Year(Max(Date)))"},_typeFact={2}, PeriodID = {"<=$(=Max(PeriodID))"} >}Followed))
# of Social Engineering emails PY	num(sum({<Year = {"$(=(max(Year)-1))"}, PeriodID = {"<=$(=Max(PeriodID))"} ,_typeFact={2} >}Followed))
Gauge: %Reported emails	(count({<[_typeFact]={2}, Reported={1}, PeriodID = {$(vMaxPeriodSE)}>}UserID)/count({<[_typeFact]={2}, Reported={1,0}, PeriodID = {$(vMaxPeriodSE)}>}UserID))*100
Last password assessment date:	Date(Max(Date_assessment))
Total Number of Passwords:	sum({<_typeFact={3},_typePassword={'SO','SSO'}, Date={"$(=Date(Max(Date_assessment)))"} >}Total_Pass)
Total Number of Complex passwords:	sum({<_typeFact={3},_typePassword={'SO','SSO'}, Date={"$(=Date(Max(Date_assessment)))"} >}Total_Pass)- sum({<_typeFact={3},_typePassword={'SO','SSO'}, Date={"$(=Date(Max(Date_assessment)))"} >}Total_Pass_Not_Compliance)
Current ratio of compliance	num((sum({<_typeFact={3},_typePassword={'SO','SSO'}, Date={"$(=Date(Max(Date_assessment)))"} >}Total_Pass) - sum({<_typeFact={3},_typePassword={'SO','SSO'}, Date={"$(=Date(Max(Date_assessment)))"} >}Total_Pass_Not_Compliance))/100, '#.##0') &':'& num(sum({<_typeFact={3},_typePassword={'SO','SSO'}, Date={"$(=Date(Max(Date_assessment)))"} >}Total_Pass)/100,'#.##0')

Compliance level	(SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vMaxPeriodPass)}, Year=, Quarter=, Month=>}Total_Pass) -SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vMaxPeriodPass)}, Year=, Quarter=, Month= >}Total_Pass_Not_Compliance)) / SUM({<_typeFact={3},_typePassword={'SO','SSO'},PeriodID = {$(vMaxPeriodPass)}, Year=, Quarter=, Month= >}Total_Pass)
Complex passwords:	(SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vMaxPeriodPass)}, Year=, Quarter=, Month=>}Total_Pass) -SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vMaxPeriodPass)}, Year=, Quarter=, Month= >}Total_Pass_Not_Compliance)) / SUM({<_typeFact={3},_typePassword={'SO','SSO'},PeriodID = {$(vMaxPeriodPass)}, Year=, Quarter=, Month= >}Total_Pass)
='Previous Period:'	(SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vPreviousPeriodPass)},Year=, Quarter=, Month= >}Total_Pass) -SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vPreviousPeriodPass)}, Year=, Quarter=, Month= >}Total_Pass_Not_Compliance)) / SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vPreviousPeriodPass)},Year=, Quarter=, Month= >}Total_Pass)
% Non crackable passwords:	(SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vMaxPeriodPass)}, Quarter=, Month=>}Total_Pass) -SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vMaxPeriodPass)}, Quarter=, Month=>}Crackable)) / SUM({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vMaxPeriodPass)}, Quarter=, Month=>}Total_Pass)
Period Avg crackable passwords:	AVG({<_typeFact={3},_typePassword={'SO','SSO'}, PeriodID = {$(vMaxPeriodPass)}, Quarter=, Month=>}Crackable)
Compliance	(SUM({<_typeFact={3},_typePassword={'SO','SSO'}, Year = {"$(=Year(Max(Date)))"} >}Total_Pass) - SUM({<_typeFact={3},_typePassword={'SO','SSO'},Year = {"$(=Year(Max(Date)))"}>}Total_Pass_Not_Compliance)) / (SUM({<_typeFact={3},_typePassword={'SO','SSO'},Year = {"$(=Year(Max(Date)))"}>}Total_Pass))

Trend	linest_m(total aggr(if((SUM({<_typeFact={3},_typePassword={‘SO’,’SSO’}, Year = {“$(=Year(Max(Date)))”} >}Total_Pass),(SUM({<_typeFact={3},_typePassword={‘SO’,’SSO’}, Year = {“$(=Year(Max(Date)))”} >}Total_Pass)),Month),Month)* only({1}Month)+linest_b(total aggr(if((SUM({<_typeFact={3},_typePassword={‘SO’,’SSO’}, Year = {“$(=Year(Max(Date)))”} >}Total_Pass),(SUM({<_typeFact={3},_typePassword={‘SO’,’SSO’}, Year = {“$(=Year(Max(Date)))”} >}Total_Pass)),Month),Month)
Table: Date of Assessment	Date_assessment
Table: Password Type	[Password Type]
Table: Avg Length	[Avg Length]
Table: Complexity	num(Complexity,'#.##0')
Table: Total # Passwords	Total_Pass
Table: # Not Compliance Pass.	Total_Pass_Not_Compliance
Table: Crackable	Crackable
Cost of Information Incidents	sum({<_typeFact={4}, Year = {"$(=Year(Max(Date)))"}, PeriodID = {"<=$(=Max(PeriodID))"} >}H_Time)*vWage
Cost of Information Incidents PY	sum({<_typeFact={4}, Year = {"$(=(max(Year)-1))"}, PeriodID = {"<=$(=Max(PeriodID))"}>}H_Time)*vWage
# of Security Incidents	num(count({<_typeFact={4}, Year = {"$(=Year(Max(Date)))"},_typeFact={4}, PeriodID = {"<=$(=Max(PeriodID))"} >}[Incident ID]))
# of Security Incidents PY	num(count({<Year = {"$(=(max(Year)-1))"}, PeriodID = {"<=$(=Max(PeriodID))"} ,_typeFact={4} >}[Incident ID]))
# of Security Improvements	num(sum({<[_typeFact]={4}, Year = {"$(=Year(Max(Date)))"},_typeFact={4}, PeriodID = {"<=$(=Max(PeriodID))"} >}Improvements))
# of Security Incidents not Solved	count({<[_typeFact]={4}, Year = {"$(=Year(Max(Date)))"}, PeriodID = {"<=$(=Max(PeriodID))"} , Status={'Work in progress'} ,_typeFact={4} >}[Incident ID])
# of Security Incidents not Solved PY	count({<Year = {"$(=(max(Year)-1))"}, PeriodID = {"<=$(=Max(PeriodID))"} , Status={'Work in progress'} ,_typeFact={4} >}[Incident ID])
=Max(Year)&' Incidents'	count({<Year = {"$(=Year(Max(Date)))"},_typeFact={4} >}[Incident ID])
=Max(Year)-1&' Incidents'	count({<Year = {"$(=(max(Year)-1))"} ,_typeFact={4} >}[Incident ID])
=Max(Year)&' Sec Improvements'	sum({<Year = {"$(=Year(Max(Date)))"},_typeFact={4} >}Improvements)
=Max(Year)-1&' Sec Improvements'	Sum({<Year = {"$(=(max(Year)-1))"} ,_typeFact={4} >}Improvements)

# of Incidents by type (Month/Incident)	num(count({<Year = {"$(=Year(Max(Date)))"},_typeFact={4}, PeriodID = {"<=$(=Max(PeriodID))"} >}[Incident ID]))
# of Assets	count({<[_typeFact]={5}, Year = {$(vMaxYearPentest)}, PeriodID={$(vMaxVulnPeriod)}, Assessment={0,1}, [BIA Level]={0,1,2,3}>} UUID_Asset_ID)
Critical Assets	count({<[_typeFact]={5}, Year = {$(vMaxYearPentest)}, PeriodID={$(vMaxVulnPeriod)}, Assessment={0,1}, [BIA Level]={3}>} UUID_Asset_ID)

% Vuln./Pentest Assessments	count({<[_typeFact]={6}, PeriodID=, Quarter=, Month=, Year = {$(vMaxYearPentest)}, Pentest={1}, [_Assessment]={1}>}UUID_Asset_ID)
	/
	count({<[_typeFact]={6}, PeriodID=, Quarter=, Month=, Year = {$(vMaxYearPentest)}, Pentest={0,1}, [_Assessment]={0,1}>}UUID_Asset_ID)

% Unpatched Vulnerabilities	count({<[_typeFact]={5}, [BIA Level]={3}, Year = {"$(=Year(Max(Date)))"}, PeriodID = {$(vMaxVulnPeriod)}, Assessment={1}, Patch={0}>}UUID_Asset_ID)
	/
	count({<[_typeFact]={5}, [BIA Level]={3}>} distinct UUID_Asset_ID)

Ratio of Vuln. Coverage	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, PeriodID ={$(vMaxVulnPeriod)}, Assessment={1}>}UUID_Asset_ID)
	/
	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, PeriodID ={$(vMaxVulnPeriod)}, Assessment={1,0}>}UUID_Asset_ID)

# Assessments by BIA Type	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, [BIA Level]={1}, Assessment={1}>}UUID_Asset_ID)
	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, [BIA Level]={2}, Assessment={1}>}UUID_Asset_ID)
	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, [BIA Level]={3}, Assessment={1}>}UUID_Asset_ID)
# Patch Coverage	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, [BIA Level]={1}, Patch={1}>}UUID_Asset_ID)
	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, [BIA Level]={2}, Patch={1}>}UUID_Asset_ID)
	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, [BIA Level]={3}, Patch={1}>}UUID_Asset_ID)
Assessments YoY Compraison	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date)))"}, [BIA Level]-={3}, Assessment={1,2,3}>}UUID_Asset_ID)

	count({<[_typeFact]={5}, Year = {"$(=Year(Max(Date))-1)"}, [BIA Level]={1,2,3}, Assessment={1}>}UUID_Asset_ID)

% Pentest Assessments	count({<[_typeFact]={6}, PeriodID=, Quarter=, Month=, Year = {$(vMaxYearPentest)}, Pentest={1}>}UUID_Asset_ID)
	/
	count({<[_typeFact]={6}, PeriodID=, Quarter=, Month=, Year = {$(vMaxYearPentest)}, Pentest={0,1}>}UUID_Asset_ID)

Gauge: Threshold status	count({<[_typeFact]={6}, Quarter=, Month=, Year = {$(vMaxYearPentest)}, Pentest={1}>}UUID_Asset_ID) / (count({<[_typeFact]={6}, Quarter=, Month=, Year = {$(vMaxYearPentest)}, Pentest={0,1}>}UUID_Asset_ID))*100
# Pentest Assessments BIA L3	count({<[_typeFact]={6}, Year = {$(vMaxYearPentest)}, Pentest={1}, [BIA Level]={3}>}UUID_Asset_ID)
Total Pentest performed YoY	count({<[_typeFact]={6}, Year=, Quarter=, Month=, Pentest={1}>}UUID_Asset_ID)
Authorized Devices	count({<[_typeFact]={7}, Year=, Month=, Quarter=, PeriodID={$(vMaxAssetPeriod)}, Authorized={1} >} distinct UUID_Asset_ID)
Unauthorized Devices	count({<[_typeFact]={7}, Year=, Month=, Quarter=, PeriodID={$(vMaxAssetPeriod)},Authorized={0}>}[IP Address])
Critical Assets (BIA Level 3)	count({<[_typeFact]={7}, [BIA Level]={3}, Year = {"$(=Year(Max(Date)))"}, PeriodID = {$(vMaxAssetPeriod)}>}UUID_Asset_ID)
Critical Assets Protected	count({<[_typeFact]={7}, [BIA Level]={3}, Patch={1}, Year = {"$(=Year(Max(Date)))"}, PeriodID = {$(vMaxAssetPeriod)}>}UUID_Asset_ID) /count({<[_typeFact]={7}, [BIA Level]={3}, Patch={1,0}, Year = {"$(=Year(Max(Date)))"}, PeriodID = {$(vMaxAssetPeriod)}>}UUID_Asset_ID)
Table1: UUID	only({<[_typeFact]={7}, Authorized={1}>}UUID_Asset_ID)
Table1: IP Address	[IP Address]
Table1: System Name	[System Name]
Table1: BIA Level	[BIA Level]
Table2: UUID	only({<[_typeFact]={7}, Authorized={0}>}UUID_Asset_ID)
Table2: IP Address	[IP Address]
Table2: System Name	[System Name]

Authorized/Unauthorized Devices	count({<[_typeFact]={7}, Year=, Month=, Quarter=, PeriodID={$(vMaxAssetPeriod)}, Authorized={1} >} distinct UUID_Asset_ID)

12.Appendix E – Code

```
SET ThousandSep='.';
SET DecimalSep=',';
SET MoneyThousandSep=',';
SET MoneyDecimalSep='.';
SET MoneyFormat='$#,##0.00;-$#,##0.00';
SET TimeFormat='h:mm:ss TT';
SET DateFormat='DD/MM/YYYY';
SET TimestampFormat='DD/MM/YYYY h:mm:ss[.fff] TT';
SET FirstWeekDay=6;
SET BrokenWeeks=1;
SET ReferenceDay=0;
SET FirstMonthOfYear=1;
SET CollationLocale='en-US';
SET CreateSearchIndexOnReload=1;
SET MonthNames='Jan;Feb;Mar;Apr;May;Jun;Jul;Aug;Sep;Oct;Nov;Dec';
SET
LongMonthNames='January;February;March;April;May;June;July;August;Septemb
er;October;November;December';
SET DayNames='Mon;Tue;Wed;Thu;Fri;Sat;Sun';
SET
LongDayNames='Monday;Tuesday;Wednesday;Thursday;Friday;Saturday;Sunday';
SET    NumericalAbbreviation='3:k;6:M;9:G;12:T;15:P;18:E;21:Z;24:Y;-3:m;-
6:μ;-9:n;-12:p;-15:f;-18:a;-21:z;-24:y';


// === Load Variables
Variables:
LOAD
     [Variable Name],
     Value,
     Comment
FROM [lib://Auxiliary/SA_Variables.xlsx]
(ooxml, embedded labels, table is Variables)
where Load = 1;

Let vNumberOfRows = NoOfRows('Variables');
For vI = 0 to (vNumberOfRows - 1)
    Let vVariable_Name = Peek('Variable Name',vI,'Value');
    Let [$(vVariable_Name)] = Peek('Value',vI,'Value');
Next
DROP Table Variables;

LET vI = Null();
LET vVariable_Name = Null();
LET vNumberOfRows = Null();


Set dataManagerTables = '','Data';
//This  block  renames  script  tables  from  non  generated  section  which
conflict with the names of managed tables

For each name in $(dataManagerTables)
    Let index = 0;
    Let currentName = name;
    Let tableNumber = TableNumber(name);
    Let matches = 0;
    Do while not IsNull(tableNumber) or (index > 0 and matches > 0)
        index = index + 1;
        currentName = name & '-' & index;
        tableNumber = TableNumber(currentName)
        matches = Match('$(currentName)', $(dataManagerTables));
    Loop
    If index > 0 then
            Rename Table '$(name)' to '$(currentName)';
    EndIf;
Next;
Set dataManagerTables = ;


Unqualify *;

__countryGeoBase:
LOAD
      ISO3Code AS [__ISO3Code],
      ISO2Code AS [__ISO2Code],
      Polygon AS [__Polygon]
FROM [lib://GEO_TABLES/countryGeo.qvd]
(qvd);

__countryCodeIsoTwo2Polygon:
MAPPING LOAD
      __ISO2Code,
      __Polygon
RESIDENT __countryGeoBase;

TAG FIELD [Country] WITH '$geoname', '$relates_Data.Country_GeoInfo';
TAG   FIELD   [Data.Country_GeoInfo]   WITH   '$geopolygon',   '$hidden',
'$relates_Country';

DROP TABLES __countryGeoBase;


// // Flag for Fact type
// @This section describes the flags used for the multiple facts available
in the General Fact Table
// 1 = Training and Awareness
// 2 = Social Engineering
// 3 - Password Complexity
// 4 - Incident
// 5 - Vulnerability
// 6 - Pentest
// 7 -Assets Inventory


//Training and awareness Data
[Facts]:
LOAD
      RowNo()
      as #factNum,
    UserID,
      [Country],
      [Job Pos],
      [Date],
    Date
      as [%DateKey],
      [Mandatory],
      [Completed],
    if([C Suite]=1,1,if([Mngmnt]=1,2,3))              as %employeeCat,
      [Score],
    if(Score<60,0,1)                                              as
_hasPassed, //<60, user failed the exam
      [C Suite],
      [Mngmnt],
      [Employee],
    FileBaseName() as Source,
    '1' as _typeFact,
      APPLYMAP( '__countryCodeIsoTwo2Polygon', UPPER([Country]), '-') AS
[Data.Country_GeoInfo]
FROM [lib://00_Stage/Training and Awareness 201*.xlsx]
(ooxml, embedded labels, table is Data);


// // The Data Set for Password complexity is aggregated by week analysis
// //Password complexity Data - System
Concatenate(Facts)
Facts:
LOAD
      RowNo() as #factNum,
    "Index",
    Pass_Category,
    if(mixmatch(Pass_Category,'SystemSO'),1,0)        as %passCat,
    //"Year",
    Date_assessment,
    (Date_assessment)                                             as
[%DateKey],
    Date_assessment
      as Date,
    "Avg Length",
    Complexity,
      //AVG complexity of all assessed passwords
    AVG_Similarity_to_user,
    Total_Pass,
      //Total number of passwords of type
    Crackable,
      //Total number of crackable passwords
    Total_Pass_Not_Compliance,
    Num(Total_Pass - Total_Pass_Not_Compliance)       as Compliance,
    FileBaseName() as Source,
    '3' as _typeFact,
    'SO' as _typePassword
FROM [lib://00_Stage/PasswordDataset.xlsx]
(ooxml, embedded labels, table is SO);


////------------- Password complexity Data - Single Sign On
Concatenate(Facts)
LOAD
      RowNo()
      as #factNum,
    "Index",
    Pass_Category,
    if(mixmatch(Pass_Category,'SSO'),2,0)             as %passCat,
    //"Year",
    Date_assessment,
    (Date_assessment)                                            as
[%DateKey],
        Date_assessment                                          as
Date,
    "Avg Length",
    Complexity,
    AVG_Similarity_to_user,
    Total_Pass,
    Crackable,
    Total_Pass_Not_Compliance,
    Num(Total_Pass - Total_Pass_Not_Compliance)       as Compliance,
    FileBaseName() as Source,
    '3' as _typeFact,
    'SSO' as _typePassword
FROM [lib://00_Stage/PasswordDataset.xlsx]
(ooxml, embedded labels, table is SSO);


////------------- Social Engineering data set:
Concatenate (Facts)     //concatenates to the data model Fact table
LOAD
      RowNo()
      as #factNum,
    "UserID",
    Date,
    Date
      as [%DateKey],
    Followed,
    Reported,
    FileBaseName()
      as Source,
    '2' as _typeFact
FROM [lib://00_Stage/SocialEngSet 20*.xlsx]
(ooxml, embedded labels, table is Assessment);



////------------- Incidents:
concatenate (Facts)
LOAD
      RowNo()

                  as #factNum,
    "Service Component WBS (aff)",
    "Incident ID",
    Status,
    Impact as %Impact,
    Urgency,
    Priority,
    Category,
    "KM number",
    "Alert Status",
    "# Reassignments",
    date(num(if((IsNum([Open                          Time])),floor([Open
Time]),date#(left(text([Open Time]),10),'DD-MM-YYYY'))))          as
%DateKey,
    date(num(if((IsNum([Open                          Time])),floor([Open
Time]),date#(left(text([Open Time]),10),'DD-MM-YYYY'))))          as
Date,
    [Open Time],
    "Reopen Time",
    if(not isnull([Resolved Time]),
            date(num(if((IsNum([Resolved          Time])),floor([Resolved
Time]),date#(left(text([Resolved Time]),10),'DD-MM-YYYY')))),
                  null()) as date_Resolved,
    [Resolved Time],
    date(num(if((IsNum([Close                        Time])),floor([Close
Time]),date#(left(text([Close Time]),10),'DD-MM-YYYY'))))   as
date_Closed,
    [Close Time],
    "Handle Time (Hours)",
    if([Handle Time (Hours)]>150,[Handle Time (Hours)]/10)
                                                                as
H_Time,
    FileBaseName()

                  as Source,
    Improvements,
     '4'

                  as _typeFact
FROM [lib://00_Stage/Incident.xlsx]
(ooxml, embedded labels, table is Incident);


//  hourly  wage  https://yalantis.com/blog/cost-services-europe-market-
research/

//------------- Vulnerabilities:
Concatenate(Facts)
LOAD
       RowNo()
      as #factNum,
     UUID_Asset_ID,
     [BIA Level],
     Assessment,
     [Date Assessment]  as Date,
     [Date Assessment]  as [%DateKey],
     Patch,
     [Date Patch],
     FileBaseName()     as Source,
     '5'                      as _typeFact

FROM [lib://00_Stage/Vulnerability_201*.xlsx]
(ooxml, embedded labels, table is Vulnerability);

//------------- Pentest
Concatenate(Facts)
LOAD
       RowNo()                as #factNum,
     UUID_Asset_ID,
     [BIA Level],
     Pentest,
     [Date_Pentest]     as Date,
     [Date_Pentest]     as [%DateKey],
      FileBaseName()    as Source,
      _Assessment,
     '6'                      as _typeFact
FROM [lib://00_Stage/Pentest_20*.xlsx]
(ooxml, embedded labels, table is Vulnerability);
```

```
//------------- Assets Inventory
Concatenate(Facts)
LOAD
       RowNo()                as #factNum,
     if(isnull(UUID_Asset_ID)                                          or
len(trim(UUID_Asset_ID))=0,'FF9999',UUID_Asset_ID) as UUID_Asset_ID,
     if(isnull(UUID_Asset_ID),'0',1) as Authorized,
     [BIA Level],
     [Date Assessment]  as Date,
     [Date Assessment]  as [%DateKey],
     [IP Address],
     Patch,
     [System Name],
     FileBaseName()     as Source,
     '7'                      as _typeFact

FROM [lib://00_Stage/Vulnerability_2019_3q.xlsx]
(ooxml, embedded labels, table is Vulnerability);



[Password Category]:
LOAD * INLINE [
    %passCat, Password Type
    1, SO
    2, SSO
];


[Employee Category]:
LOAD * INLINE [
    %employeeCat, Employee Category
    1, C Suite
    2, Management
    3, Regular Employee
];


[Incident Type]:
LOAD
    Impact as %Impact,
    "Incident Type",
    "Priority Level"
FROM [lib://00_Stage/Incident.xlsx]
(ooxml, embedded labels, table is [Incident Type])
where Impact <> 0;


Temp_Calendar_Range:
LOAD
      Num(Min(Date))                as MinDate,
      Num(Max(Date))                as MaxDate
RESIDENT [Facts];

LET vMinDateTemp = Peek('MinDate', 0, 'Temp_Calendar_Range');
LET vMaxDateTemp = Peek('MaxDate', 0, 'Temp_Calendar_Range');

DROP TABLE Temp_Calendar_Range;

[Master Calendar]:
Load *,
       AutoNumber(Year & Quarter, 'QuarterID')        as [QuarterID],
       AutoNumber(Year(%DateKey)&Month(%DateKey), 'PeriodID')
      as [PeriodID]
       ;

LOAD DISTINCT
      Temp_Date
      as [%DateKey],
      Year(Temp_Date)
      as [Year],
      Month(Temp_Date)                                            as
[Month],
      Day(Temp_Date)
      as [Day],
    WeekDay(Temp_Date)                                            as
WeekDay,
    Week(Temp_Date)                                               as
Week,
      num(Month(Temp_Date))                                       as
[MonthNum],
      Date(Temp_Date, 'DD-MM-YYYY')                         as   [Year   -
Month],
      'Q' & Ceil(Month(Temp_Date) / 3)                      as
[Quarter]//,
      ;
LOAD DISTINCT
      Date($(vMinDateTemp) + IterNo() - 1)   as Temp_Date
AUTOGENERATE (1)
WHILE $(vMinDateTemp) + IterNo() - 1 <= $(vMaxDateTemp);

LET vMinDateTemp = Null();
LET vMaxDateTemp = Null();


Drop Fields
      [C Suite],
      [Mngmnt],
      [Employee]
  ;
```