NIST Special Publication 800
NIST SP 800-55v1 ipd

Measurement Guide for Information Security

Volume 1 — Identifying and Selecting Measures

Initial Public Draft

Katherine Schroeder
Hung Trinh
Victoria Yan Pillitteri

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.800-55v1.ipd

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Special Publication 800
NIST SP 800-55v1 ipd

Measurement Guide for Information Security

Volume 1 — Identifying and Selecting Measures

Initial Public Draft

Katherine Schroeder
Hung Trinh
Victoria Yan Pillitteri
Computer Security Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.800-55v1.ipd

January 2024

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

Authority

NIST Technical Series Policies

Publication History

Approved by the NIST Editorial Review Board on YYYY-MM-DD [Will be added in the final publication]
Supersedes NIST Series XXX (Month Year) DOI [Will be added in the final publication]

How to Cite this NIST Technical Series Publication:

Schroeder K, Trinh H, Pillitteri V (2024) Measurement Guide for Information Security: Volume 1 — Identifying and Selecting Measures. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) NIST SP 800-55v1 ipd. https://doi.org/10.6028/NIST.SP.800-55v1.ipd

Author ORCID iDs

Katherine Schroeder: 0000-0002-4129-9243
Hung Trinh: 0000-0002-3323-0836
Victoria Yan Pillitteri: 0000-0002-7446-7506

Public Comment Period
January 17, 2024 – March 18, 2024

Submit Comments
cyber-measures@list.nist.gov

National Institute of Standards and Technology
Attn: Computer Security Division, Information Technology Laboratory
100 Bureau Drive (Mail Stop 8930) Gaithersburg, MD 20899-8930

All comments are subject to release under the Freedom of Information Act (FOIA).

Abstract

This document provides guidance on how an organization can develop information security measures to identify the adequacy of in-place security policies, procedures, and controls. It explains the measures prioritization process and how to evaluate measures.

Keywords

assessment; information security; measurement; measures; metrics; performance; qualitative; quantitative; reports; security controls.

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at the National Institute of Standards and Technology (NIST) promotes the U.S. economy and public welfare by providing technical leadership for the Nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analyses to advance the development and productive use of information technology. ITL's responsibilities include the development of management, administrative, technical, and physical standards and guidelines for the cost-effective security and privacy of other than national security-related information in federal information systems. The Special Publication 800-series reports on ITL's research, guidelines, and outreach efforts in information system security, and its collaborative activities with industry, government, and academic organizations.

Audience

This guide is written primarily for users with responsibilities or interest in information security measurement and assessment. Government and industry can use the concepts, processes, and candidate measures presented in this guide.

Call for Patent Claims

This public review includes a call for information on essential patent claims (claims whose use would be required for compliance with the guidance or requirements in this Information Technology Laboratory (ITL) draft publication). Such guidance and/or requirements may be directly stated in this ITL Publication or by reference to another publication. This call also includes disclosure, where known, of the existence of pending U.S. or foreign patent applications relating to this ITL draft publication and of any relevant unexpired U.S. or foreign patents.

ITL may require from the patent holder, or a party authorized to make assurances on its behalf, in written or electronic form, either:

a) assurance in the form of a general disclaimer to the effect that such party does not hold and does not currently intend holding any essential patent claim(s); or

b) assurance that a license to such essential patent claim(s) will be made available to applicants desiring to utilize the license for the purpose of complying with the guidance or requirements in this ITL draft publication either:

    i. under reasonable terms and conditions that are demonstrably free of any unfair discrimination; or

    ii. without compensation and under reasonable terms and conditions that are demonstrably free of any unfair discrimination.

Such assurance shall indicate that the patent holder (or third party authorized to make assurances on its behalf) will include in any documents transferring ownership of patents subject to the assurance, provisions sufficient to ensure that the commitments in the assurance are binding on the transferee, and that the transferee will similarly include appropriate provisions in the event of future transfers with the goal of binding each successor-in-interest.

The assurance shall also indicate that it is intended to be binding on successors-in-interest regardless of whether such provisions are included in the relevant transfer documents.

Such statements should be addressed to: cyber-measures@list.nist.gov.

Note to Reviewers

The initial public drafts (ipd) of NIST Special Publication (SP) 800-55, *Measurement Guide for Information Security, Volume 1 – Identifying and Selecting Measures* and *Volume 2 – Developing an Information Security Measurement Program* are available for comment after extensive research, development, and customer engagement.

In response to the feedback from the pre-draft call for comment and initial working draft (annotated outline), NIST continued to refine the publications by organizing the guidance into two volumes and developing more actionable and focused guidance in each.

- Volume 1 – *Identifying and Selecting Measures* – is a flexible approach to the development, selection, and prioritization of information security measures. This volume explores both quantitative and qualitative assessment and provides basic guidance on data analysis techniques as well as impact and likelihood modeling.
- Volume 2 – *Developing an Information Security Measurement Program* - is a methodology for developing and implementing a structure for an information security measurement program.

Reviewers are encouraged to comment on all or parts of draft NIST SP 800-55 *Measurement Guide for Information Security, Volume 1 – Identifying and Selecting Measures* and *Volume 2 – Developing an Information Security Measurement Program*. NIST request comments be submitted to cyber-measures@list.nist.gov by 11:59 PM Eastern Time (ET) on March 18, 2024. Commenters are encouraged to use the comment template provided with the document announcement.

Table of Contents

List of Tables

List of Figures

1. Introduction

Information security measurement enables organizations to describe and quantify information security, allocate finite resources, and make informed and data-driven decisions. However, organizations first need to know what policies, procedures, and controls they have in place at any given time; whether those countermeasures are working effectively and efficiently; and how the organization and its risks are impacted. By developing and monitoring measurements that evaluate what an organization has in place for information security risk management and how well those efforts are working, an organization can better address their goals and direct resources.

1.1. Purpose and Scope

NIST Special Publication (SP) 800-55v1, r2 is a flexible guide to the development and selection of information security measures at the organization, mission/business, and system levels to identify the success of in-place policies, procedures, and controls.[1] This document expands on previous NIST work on information security measures and measurements by focusing on quantitative assessment[2] and addressing organizational or program maturity.

The Measurement Guide for Information Security, Volume 2 – Program, provides a methodology for implementing an information security measurement program. Additionally, while many of the principles of information security measurement may apply to privacy, privacy is out of scope for this document.

1.2. Relationship to Other NIST Publications

This document is intended to provide considerations for measuring the information security program activities described in other NIST publications, including:

- SP 800-137A, *Assessing Information Security Continuous Monitoring Programs* [14]
- *Framework for Improving Critical Infrastructure Cybersecurity, Version* 2.0 (NIST Cybersecurity Framework) [1]
- SP 800-30r1, *Guide for Conducting Risk Assessments* [9]
- SP 800-37r2, *Risk Management Framework for Information Security Systems and Organizations: A System Life Cycle Approach for Security and Privacy* [10]
- SP 800-161r1, *Cybersecurity Supply Chain Risk Management Practices for Systems and Organizations* [17]
- NIST Engineering Handbook [18]

[1] This document uses the term *controls* to broadly describe identified countermeasures for managing information security risks. It is intended to be framework- and standard-agnostic and can also apply to other existing models or frameworks.
[2] SP 800-55 uses the terms *quantitative assessment* and *measurement* synonymously. Refer to Sec. 1.4, Document Terminology, for additional information.

- NIST Internal Report (IR) 8286, *Identifying and Estimating Cybersecurity Risk for Enterprise Risk Management (ERM)* [4]

1.3. Document Organization

The remaining sections of this document discuss the following:

- Section 2, Information Security Measurement Development Process
- Section 3, Measurement Development and Selection
- Appendix A, Glossary
- Appendix B, Data Analysis Dictionary
- Appendix C, Likelihood and Impact Models
- Appendix D, Change Log

1.4. Document Terminology

In the context of this document, the follow terms are defined as follows:

- **Assessment:** The action of evaluating, estimating, or judging against defined criteria. Different types of assessment (qualitative, quantitative, and semi-quantitative) are used to assess risk. Some types of assessment yield measures.
- **Assessment result:** The output or outcome of an assessment.
- **Information security[3]:** The protection of information and systems from unauthorized access, use, disclosure, disruption, modification, or destruction to provide confidentiality, integrity, and availability. [2]
- **Measurement:** The process of obtaining quantitative values using quantitative methods.
- **Measures:** Quantifiable and objective values that result from measurement.
- **Metrics:** Measures and assessment results designed to track progress, facilitate decision-making, and improve performance with respect to a set target.
- **Qualitative assessment:** The use of a set of methods, principles, or rules for assessing risk based on nonnumerical categories or levels. [9]
- **Quantitative assessment:** The use of a set of methods, principles, or rules for assessing risks based on the use of numbers where the meanings and proportionality of values are maintained inside and outside of the context of the assessment. [9]
- **Semi-quantitative assessment:** The use of a set of methods, principles, or rules for assessing risk based on bins, scales, or representative numbers whose values and meanings are not maintained in other contexts. [9]

[3] The term *information security* can be used interchangeably with *cybersecurity*.

This document discusses concepts that are similar to the Stevens Scale of Measurement, as shown in ***Table 1****, but takes a different view on what is and is not a measurement. For the purposes of this document, a nominal scale is considered a form of data gathering, and an ordinal scale is considered a ranking system. Both interval and ratio scales use variables that represent true numbers and can be used in a quantitative assessment, so they are considered measurement [19].*

Table 1. Stevens Scale of Measurement

Scale Level	Definition
Nominal	A nominal scale only looks at classification or identification. Nominal scales are used in surveys and in dealings with either non-numeric variables or numbers that do not have an assigned value. The data collected from a nominal scale can be used for counting, mode, or correlation contingency matrices.
Ordinal	An ordinal scale is similar to a nominal scale in that it primarily uses non-numeric values or numbers that are meant to show ranking. Related statistics include medians and percentiles.
Interval	An interval scale is used when measuring variables with equal intervals between values. When using an interval scale, there is no true zero. Examples of the use of interval scales are temperature or time scales. Interval data allows for quantitative analysis, such as descriptive statistics like frequency, averages, position, and dispersion. Interval statistics include mean, standard deviation, and rank-order correlation.
Ratio	Ratio scales allow for the categorization and ranking of data, similar to an interval scale, but with a true zero and no negative values. Ratio scales allow for numbers to be used for addition, subtraction, multiplication, and division.

2. Fundamentals

The terms *measurement* and *assessment* are often used interchangeably in the information security field. This document provides a lexicon for key terminology and an overview of foundational concepts to those looking to measure and assess information security risk and clarifies the distinction between measurement and assessment. As described in Sec. 1.4, assessment refers to the process of evaluating, estimating, or judging against defined criteria, and measurement is the process of obtaining quantitative values. Hence, assessment is a broader concept that also includes measurement.

Organizations perform multiple kinds of assessment when evaluating information security risk, such as risk assessments, program assessments, and control assessments. Risk assessments are used to identify the risks that an organization faces and can support decision-making [9]. Program-level assessments are used for decision-making about the strategies, policies, procedures, and operations that determine the security posture of an information security program. In control assessments, organizations evaluate whether specific controls are performing the way they were intended and achieving the desired results. Both program assessments and control assessments are in and of themselves a form of risk assessment and provide a different lens for viewing information security risk. SP 800-55 is intentionally agnostic on specific risk assessment models. However, many identify threat, likelihood, vulnerability, and impact as areas to assess.[4]

2.1. Types of Assessment

There are three types of assessment:

1. *Qualitative assessments* use non-numerical values or categories, such as high, medium, and low or heat maps.
2. *Semi-quantitative assessments* use numbers, but those numbers do not maintain their value outside of the assessment context. This is commonly seen in models that use number rankings to show a level of organizational integration. While the assessment may say that the organization is at “level 3,” that “3” represents a set of qualities rather than a numerical value.
3. With *quantitative assessments*, any numbers used retain their value outside of the context. For example, 98 % of authorized accounts belong to current employees, and 2 % belong to former employees. Here, the values “98 %” and “2 %” stay the same regardless of the context. Since measurement is the process of obtaining quantifiable values using *quantifiable assessment methods*, measures are *quantitative assessment results*.

Quantitative assessments (i.e., measurements) can provide objective data that allows for tracking and shows changes. However, they can be more difficult to produce since they require more data and resources than qualitative assessments. In contrast, qualitative assessments

[4] For additional information about risk assessment models, see [9].

may be more commonly used and easier to conduct, but their results can also be subjective and require everyone to have an equal understanding of the scale used.

Organizations will first consider their motivations for measuring information security risks before determining whether a quantitative or qualitative assessment is appropriate. For example, an organization motivated primarily by compliance with an industry certification or international standard has different measurement needs than an organization motivated by cost reduction. An organization could have multiple, competing motivations that drive the identification and selection of measures.

Some organizational motivations may benefit from quantitative assessments, such as trying to determine whether the organization is patching known vulnerabilities in an acceptable amount of time. Knowing the ***mean time to remediate a vulnerability*** *provides more precise insight into patching efficiency than simply knowing the number of vulnerabilities patched in a year. Because the question of* ***mean time to remediate a vulnerability*** *deals in non-zero numbers that are attainable to gather, a measurement can be taken, and a mathematically derived answer can be given.*

When real and attainable numbers based on gathered data can be found and analyzed, a quantitative assessment may be the appropriate action. If there are proposed questions that do not have measurable numbers attached to them but still need to be addressed, a qualitative assessment may be the best option.

Commonly used qualitative methods include color scales that represent risk levels or number scales that show rankings. For the purposes of this document, qualitative and semi-quantitative assessments are not considered measurement, and the values produced by these types of assessments are not considered measures. Most organizations will use a mixture of quantitative, semi-quantitative, and qualitative assessments. Ultimately, some or all the assessment results will be used to determine success.

In addition to measurement, organizations also utilize *metrics* to track progress, facilitate decision-making, and improve performance. Information gained from measurement may be used to identify and define new metrics. Metrics can be applied at the system level, program level,[5] and organization level. System-level metrics, such as the frequency of third-party access to a system or the number of communication ports open, can facilitate tactical decision-making and support program-level metrics. Program-level metrics, such as the number of security incidents in a year or the cost per incident, may be helpful when making organizational strategic decisions. Both system- and program-level metrics can also support risk management-informed decision-making.

[5] SP 800-39, *Managing Information Security Risk: Organization, Mission, and Information System View*, includes a model of multi-level risk management for the integration of risk management across the organization. In this model, three levels are identified to address risk: (i) the organization level, (ii) the missions/business process level, and (iii) and the system level. For the purpose of this document, the program-level can be synonymous with the mission/business process-level and/or the organization-level.

2.2. Benefits of Using Measures

Developing and establishing measures to capture and provide meaningful data at all levels of an organization requires careful consideration. Meaningful measures take organizational information security goals and objectives into account and are obtainable, repeatable, and feasible to measure. Information security measurement enables organizations to quantify improvements or gaps in securing systems and demonstrate quantifiable progress in accomplishing strategic goals and objectives. Well-designed measurements can provide information on the implementation, effectiveness, efficiency, and business impacts of controls, such as the results of information security activities, events (e.g., incident data, revenue lost to cyber attacks), and information security investments.

Measurement also provides data that can enable an organization to examine the impacts of implementing information security programs, specific controls, and associated policies and procedures. Such data is integral when making risk-based decisions, weighing performance against designated metrics, and demonstrating compliance. Measurement can also increase accountability by providing data that can facilitate the identification of the personnel responsible for controls implemented within specific organizational components or systems and support an environment that allows for continuous analysis and improvement.

2.3. Measurement and Quantitative Assessment

Measures are numerically expressed data that are gathered through the process of measurement.[6] Measures can be derived from any operations or systems that can be measured with numbers. Quantitative assessments judge measures data against a set criteria or target and can be used to analyze information security risks using frequency, rates, financial impacts, and other numeric indicators.

Using quantitative assessments requires a knowledge of measurement techniques and data analysis processes. One challenge of measurement is using the right measures and quantity of measures to perform useful analysis. A single measure alone may not provide sufficient data to make risk-based decisions, but organizations may also have restraints on resources that prevent them from employing and analyzing every potential measure. An organization finds the number of measures and depth of analysis that work best for their needs.

The ability to measure information security risk relies on data availability. Methods for collecting information security data may include experimentation, observation, or sampling. The NIST Engineering Statistics Handbook [18] offers detailed information on choosing a sampling scheme, including the following methods:

- *Experimentation* is a systematic approach to testing new ideas, methods, or activities that applies principles and techniques at the data collection stage to ensure the generation of valid, defensible, and supportable conclusions. A recognizable use of experimentation to collect information security data is a phishing test, which is a form of

[6] As described in Section 1.4, *measures* and *quantitative assessment results* can be used synonymously, as can the terms *measurement* and *quantitative assessment*.

internal security testing where organizations send fake phishing emails to determine which users respond to it. The rates of success are then judged against set criteria.

- *Observational data* refers to data captures through the observation of an activity or behavior without the direct involvement of the subject. Observational data is often gathered as part of routine information security operations, such as log management tools that are used to collect and analyze network activities. Data from these logs are observational and can be used for further analysis.
- *Sampling* is the process of taking samples of something for the purpose of analysis. Sampling may be used when continuous observation and passive data collection are not an option or when *random*, *stratified*, or *systematic sampling* may be preferred. *Random sampling* is a method of sampling where each sample has an equal chance of selection in hopes of gathering an unbiased representation. *Stratified sampling* is the process of segmenting a population across levels of some factors to minimize variability within those segments (e.g., taking a sample from a terminal in each department of an organization). *Systematic sampling* is a method of sampling where samples are taken at a regular interval (e.g., once an hour or from every tenth user).

Once the data from measurement is procured, the outputs of quantitative analysis can be used in a quantitative assessment to determine whether the organization is meeting its information security goals and support risk-based decision-making. Data analysis methods[7] are largely based off of the type of questions that the organization is asking about their information security risks, program, and controls. The NIST Engineering Statistics Handbook [18] identifies three popular approaches to data analysis:

1. Classical — The classical data analysis approach is when data collection is directly followed by modeling, and the analysis, estimation, and testing that come after focus on the parameters of that model. Classical data analysis includes deterministic and probabilistic models, such as regression and the analysis of variance (ANOVA).
2. Exploratory — Exploratory data analysis begins by inferring what model would be appropriate before trying different analytic models. Identifying patterns in the data may give insight as to what models would produce the most useful information. Some common exploratory data analysis graphical techniques include standard deviation plots and histograms.
3. Bayesian — Bayesian methodology consists of formally combining both the prior distribution of the parameters and the collected data to jointly make inferences and/or test assumptions about the model of parameters. Bayesian methods can be used for expected range setting and predictive models.

Table 2 shows examples of quantitative analysis across risk assessment, program-level assessment, and control-level assessment.

[7] Appendix C describes additional examples of quantitative data analysis methods.

Table 2. Data analysis examples

Type of Assessment	Approach	Example
Risk Assessment	Classical (Value at Risk)	An organization conducting a risk assessment will likely consider their value at risk (VaR) if they were to suffer an adverse information security event. The organization may look at potential losses from downtime, the cost of repairing the environment, or reputational damage.
Risk Assessment	Bayesian	The Bayesian method looks at prior distribution, collected data, and set parameters to make inferences about future outcomes. Using data from SP 800-53 control RA-3(4), Predictive Cyber Analytics, as part of a risk assessment, the inferences found through the Bayesian method allow organizations to make risk-based decisions based on the likelihood of future events.
Program-Level Assessment	Classical (Mean)	At the program level, an organization may choose to identify the mean time it takes to complete an action. For example, using SP 800-53 control PM-22, Personally Identifiable Information Quality Management, the mean time to correct or delete inaccurate or outdated personally identifiable information is measured. The organization may also consider the variance in that data from year to year or see whether certain individuals are addressing that personally identifiable information at different rates.
Program-Level Assessment	Exploratory Data Analysis (Scatter Plot)	An organization may want to use a scatter plot as part of a program-level assessment to reveal relationships or associations between two variables. Using data collected as part of SP 800-53 control PM-31, Continuous Monitoring Strategy, one can examine linear relationships shown in a scatter plot of historical data. The scatter plot can reveal outliers or information about what typical uses of a system.
Program-Level Assessment	Bayesian	The Bayesian method can be used to influence programmatic decisions around continuous improvement. For example, using SP 800-53 control PM-6, Measures of Performance, and the Bayesian method on prior historical data, an organization can determine what future data may look like. This information on future outcomes can be used to set the expected results of information security performance.
Control Assessment	Classical (Linear Regression)	At the control level, an organization may have implemented continuous monitoring (i.e., SP 800-53, control CA-7) of a specific system-level metric. The data provided by the continuous monitoring of a system can be used in linear regression to learn what “normal” looks like for that system,

Type of Assessment	Approach	Example
		which in turn allows the organization to identify deviations from that "normal." This is a foundational piece of the information security measurement and assessment process.
Control Assessment	Exploratory Data Analysis	At the control level, a multi-factor/comparative box plot could be used to compare the key characteristics or unusual data in a data set monitoring a control.
Control Assessment	Bayesian	The Bayesian method may be used to make decisions about the frequency of equipment maintenance using SP 800-53 control MA-6(2), Timely Maintenance | Predictive Maintenance and historical data about organizational equipment.

Organizations that are early in the process of assessing their information security risks, program, or systems may rely heavily on qualitative assessments that present non-numerical information in place of measurement. These non-numerical methods can help show context, examine labels, and look at behavior. A prominent example of qualitative assessment featured in many information security measurement programs is the risk matrix — a table that uses colored rating scales to show the impact and likelihood of various risks. As organizations gain the ability to record and track information security data, they are able to move away from the subjectivity of qualitative assessments and toward the increased precision and reduced bias of quantitative assessments.

2.4. Measurement Considerations

Because measurement can involve large amounts of data, having a plan for data handling is critical to ensure that factors such as documentation, data management, data quality, and uncertainty are all considered. An organized and repeatable process that allows for the consistent assessment of collected data provides much-needed context for measurements.

Information security measurements can be scoped to a variety of environments and needs. Assets, controls, vulnerabilities, and security incidents can all be measured. Measures can be applied to organizational units, sites, or other constructs. Organizations will carefully define the scope of measures based on specific stakeholder needs, strategic goals and objectives, operating environments, risk priorities, and resources.

Information security measures can be applied at the system level to provide quantifiable data regarding the implementation, effectiveness, efficiency, and impact of required or desired security controls. System-level measures can be used to determine the system's security posture, demonstrate compliance with organizational requirements, and identify areas of improvement.

Measurements can be used to monitor organizational information security activities at the program and organization levels. These measurements may be derived by aggregating multiple system-level measures or developed by using the entire enterprise as the scope. Organization-level measurements require that the processes on which the measures depend are consistent, repeatable, and ensure the availability of data across the organization.

Perfectly measuring information security is challenging due to the gap between mathematical models and practical implementations [21]. Instead, experimenting as possible with relative metrics, models, and approaches over time is the best way to identify the most effective performance indicators.

2.4.1. Measures Documentation

Organizations document their measures in a standard format to ensure the repeatability of measures development, collection, and reporting activities. By keeping a consistent record of what is being measured, where the data comes from, what formulas and calculations are being used, and who interacts with the data, it becomes easier to trace data and ensure continuity of the process.

Organizations can tailor their standard format to their unique environments and requirements based on internal practices and procedures. However, the following fields offer a common starting point:

- **Unique ID:** A unique identifier for tracking and sorting. The unique identifier can use an organization-specific naming convention or directly reference another source.
- **Goal:** Statement of strategic and/or information security goals to guide control implementation for system-level control measures as well as higher-level measures. These goals are usually articulated in strategic and performance plans. When possible, include both the organization-level goal and the specific information security goal extracted from organization documentation, or identify an information security goal that would contribute to the accomplishment of the selected strategic goal.
- **Measure:** Statement of measurement. Use a numeric statement that begins with the word “percentage,” number,” “frequency,” “average,” or other similar term.
- **Type:** Statement of whether this is a record of implementation or a measure of effectiveness, efficiency, or impact.
- **Formula:** Calculation that results in a numeric expression of a measure. The organization may also note the information gathered in an implementation survey.
- **Target:** Threshold for a satisfactory rating for the measure (e.g., a milestone completion or statistical measure) that can be expressed in percentages, time, currency, or other unit of measurement. The target may be tied to a required completion timeframe. It may also be useful to select and record final and interim targets to track progress toward a stated goal.

- **Implementation evidence:** Evidence used to compute the measure, validate that the activity is performed, and identify probable causes of unsatisfactory results for a specific measure.
    - For manual data collection, identify questions and data elements that would provide the data inputs necessary to calculate the measure's formula, qualify the measure for acceptance, and validate the information provided.
    - For automated data collection, identify data elements that would be required for the formula, qualify the measure for acceptance, and validate the information provided.
- **Frequency:** How often the data is collected, analyzed, and reported. Select the frequency of data collection based on a rate of change that is being evaluated. Select the frequency of data reporting based on external reporting requirements and internal customer preferences.
- **Responsible parties:** Key stakeholders, such as:
    - Information owner — Identify the organizational component and the individual who owns the required information.
    - Information collector — Identify the organizational component and the individual responsible for collecting the data.[8]
    - Information customer — Identify the organizational component and the individual who will revive the data.
- **Data source:** Location of the data to be used in calculating the measure, including databases, tracking tools, logs, organizations, and specific roles within the organization that can provide the required information.
- **Reporting format:** Indication of how the measure will be reported, such as a pie chart, line chart, bar graph, or other format. It may also be beneficial to include a sample.

2.4.2. Data Management

Although substantial amounts of information security data may be collected, not all data will be useful for the information security measurement program at any given point in time. Any data collected specifically for information security measures are as nonintrusive as possible and of maximum usefulness to ensure that available resources are primarily used to correct problems rather than collect data.

The information in information security data repositories represents a significant collection of operational and vulnerability data. Due to the sensitivity of this data, information security

[8] When possible, the information collector will be a different individual or even a representative of a different organizational unit than the information owner to avoid the possibility of a conflict of interest and ensure separation of duties, though this may not be feasible for smaller organizations.

performance measurement data repositories are protected in accordance with applicable laws, regulations, policies, and procedures.

2.4.3. Data Quality

Data collection methods and the data repositories used for measures data collection and reporting (either directly or as data sources) are clearly defined to ascertain the quality and validity of the data. This also helps ensure that testing is repeatable and can show changes over time.

Data validity is suspect if the primary data source is an incident-reporting database that only stores information reported by a few organizational elements or if reporting processes between organizations are inconsistent. The importance of standardizing reporting processes cannot be overemphasized. When organizations are developing and implementing processes that may serve as inputs into an information security measurement program, while ensuring that data gathering and reporting are clearly defined to facilitate valid data collection. Having a validation process in place to check the integrity, accuracy, and structure of the data provides a way to address potential errors before any analysis is done. By setting a standard process to validate data, an organization can have a repeatable way to look at the data and ensure its quality.

2.4.4. Uncertainty and Errors

Even when measurements are intended to be precise and accurate, random and systemic errors can still occur. While there is no guaranteed way to measure uncertainty in all measurements, statistical information calculated from the data (e.g., standard deviation, standard error of mean, and confidence intervals) can provide more insight.

Uncertainty can be reduced by using data cleaning methods, such as validation, normalization, transformation, and imputation, as shown in **Table 3**.

Table 3. Data cleaning methods for reducing uncertainty

Data Cleaning Method	Definition
Data Validation	The process of determining that collected data is acceptable according to a predefined set of tests [15]
Normalization	The conversion of information into consistent representations and categorizations [4]
Transformation	The conversion of data from one state or format into another state or format
Imputation	The replacement of unknown, unmeasured, or missing data with a particular value. The simplest form of imputation is to replace all missing values with the average of that variable [18].

In addition to making the data itself more useable, data analysis methods (e.g., sensitivity analysis and Monte Carlo analysis) can address uncertainty within the data. Organizations often make quantitative projections using statistical methods, such as regression, time series analysis, and machine learning methods. When looking at projections, it is helpful to consider that future events and other unknown factors can cause unforeseen changes.

2.5. Metrics

Metrics are designed to track progress, facilitate decision-making, and improve performance with respect to a set target. Metrics leverage measures and assessment results to provide insight into how well an organization is performing at the program or system level and whether the organization is reducing their information security risk. As with measures, the characteristics of meaningful metrics include the value being objective, accurate, precise, tied to a fixed reference or point in time, replicable, and comparable to previous measurements. Metrics are set with organizational goals in mind and drive subsequent assessments whose results then inform the metrics going forward. Figure 1 shows a notional process for the definition, collection, and analysis of metrics.

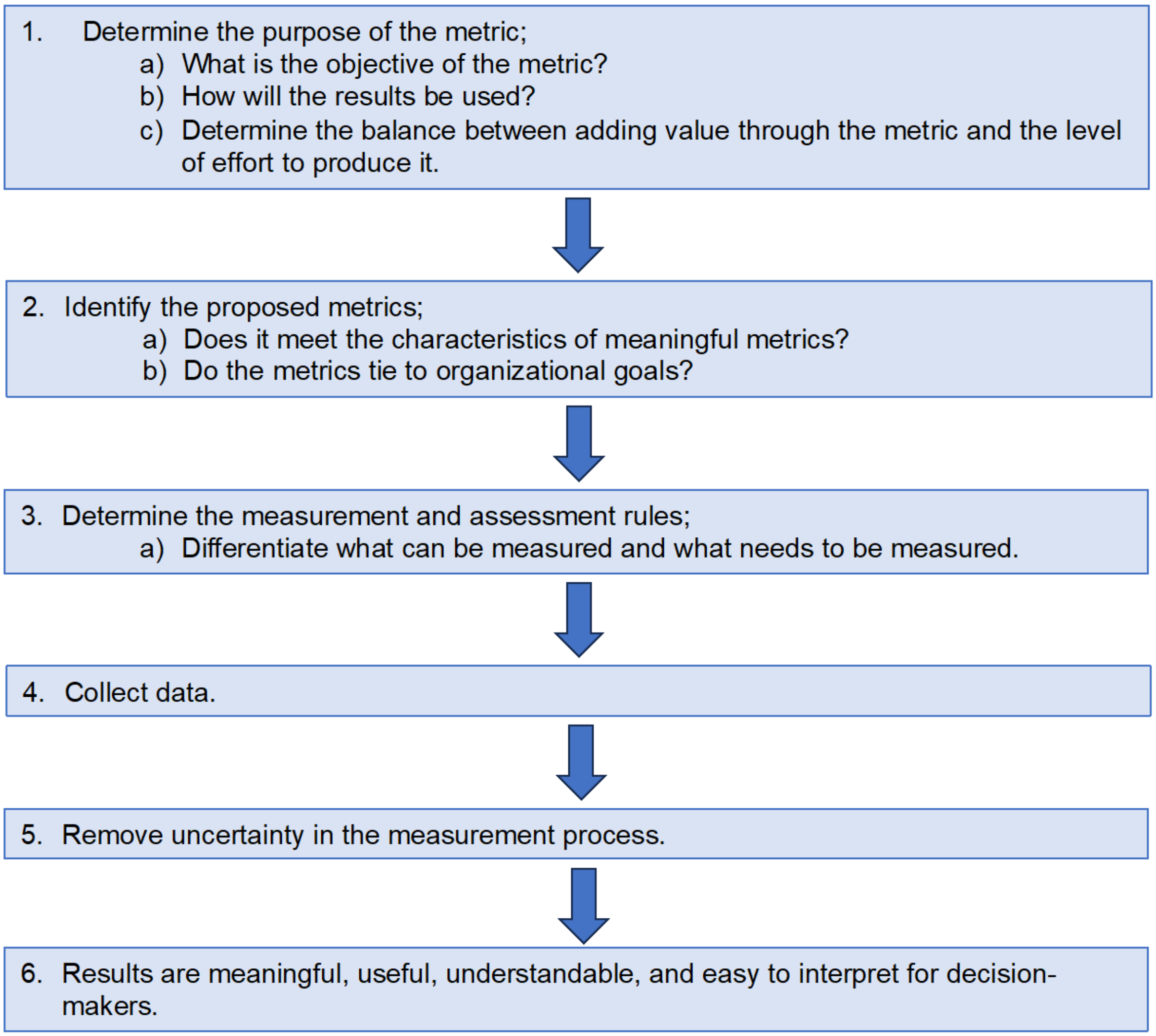

Fig. 1. Notional process for the definition, collection, and analysis of metrics

When selecting measurements and metrics to focus on, it is helpful to know why the measurements are being taken and their purpose. It is important that the chosen metrics tell a meaningful story about organization-, program-, or system-level information security. To do so, metrics are designed to be unambiguous so that their purpose and output can be more easily understood. For example, when evaluating cybersecurity awareness training, consider completion rates and the results of review quizzes instead of marking participation as “low, medium, or high.”

By keeping metrics consistent over time, organizations can evaluate long-term trends and expected ranges. A new metric may provide important insight, but tracking the measurements related to metrics over a continuous period (e.g., quarter to quarter, year to year) will give more information about the success of organization-, program-, and system-level information security plans, policies, procedures, and goals. Some metrics may be gathered because of outside guidance or regulations.

Key risk indicators (KRIs) and key performance indicators (KPIs) are examples of metrics, though not all metrics fall into these categories. Organizations may find a wide variety of metrics fit their needs. For example, appropriate measures at the organization level may include the cost per security incident as part of the budget allocation process, whereas measurements at the system level may include the frequency of virus scans across individual systems.

3. Selecting and Prioritizing Measures

Developing and selecting information security measures consists of four major activities:

1. Identifying and defining the current information security program
2. Developing specific measures to gauge the implementation, effectiveness, efficiency, and impact of security controls
3. Prioritizing measures based on organizational needs
4. Evaluating collected measures data

3.1. Identification and Definition

This document focuses on the development of measures related to information security risk management, which is part of a larger implementation process of information security measurement.[9] The identification and definition of the existing information security program are important to the development of measures.

Identification and definition include:

- **Stakeholders and interests:** Identifying relevant stakeholders and their interests in information security measurement
- **Goals and objectives:** Identifying and documenting security goals and objectives that will guide control implementation
- **Information security policies, guidelines, and procedures review:** Examining existing organization-specific policies, guidelines, and procedures related to information security
- **Information security implementation review:** Reviewing any existing measures and data repositories that can be used to derive measures data

3.2. Types of Measures

Knowing what controls are implemented in an organization is foundational to quantitative assessment. A complete understanding of the system- and program-level controls that need to be tracked are needed before an organization can evaluate what kinds of measurements to take or the process of prioritizing potential measures. This creates a structure for determining what measurements need to be taken and what metrics are used for evaluation.

There are four types of measures/assessment results:

1. Implementation
2. Effectiveness
3. Efficiency

[9] Refer to SP 800-55 Volume 2 for more information.

4. Impact

3.2.1. Implementation Measures

Implementation measures demonstrate the progress of specific controls. Monitoring implementation may include assessment results, such as a tally of known systems or a binary "yes/no" about which systems have up-to-date patches.[10] Implementation measures look at quantitative outputs and are usually demonstrated in percentages. Examples of implementation measures related to information security programs include the percentage of systems with approved system security plans and the percentage of systems with password policies that are configured as required. Implementation measures can also examine system-level areas, such as the percentage of servers in a system with a standard configuration.

By gathering this data, an organization can understand how its goals are being implemented and what tasks still need to be accomplished. Organizations never fully retire implementation measures because they are a record of what exists and what needs improvement. However, once implementation measures are completed, the emphasis and resources of the measurement program shift away from implementation to include effectiveness, efficiency, and impact measures.

3.2.2. Effectiveness Measures

Effectiveness measures evaluate how well implementation processes and controls are working and whether they are meeting the desired outcome. An effectiveness assessment can either concentrate on the evidence and results of a quantitative analysis of measures or be applied in a qualitative "yes/no" paradigm. Effectiveness measures may require multiple data points that quantify the degree to which information controls are implemented and their effects on the organization's information security posture.

3.2.3. Efficiency Measures

Efficiency measures examine the timeliness of controls by determining the speed at which they give useful feedback and how quickly those issues are addressed. An efficiency assessment concentrates on the evidence and results of quantitative measures analysis.

3.2.4. Impact Measures

Impact measures articulate the impact of information security on an organization's unique mission, goals, and objectives by quantifying the following:

- Cost savings produced by the information security program or through costs incurred from addressing information security events

[10] Records of these essential implementation assessment results are foundational to information security measurement and are addressed in SP 800-55 Volume 2.

- Business value gained or lost
- The degree of public trust gained or maintained by the information security program
- Other mission-related impacts of information security

These measures combine the results of control implementation with a variety of information about resources. They can provide the most direct insight into the value of information security to the organization and are the ones that executives seek.

3.2.5. Comparing Measures and Assessment Results

Qualitative and semi-quantitative assessments may also be useful or even necessary to assess implementation, effectiveness, efficiency, and impact, as shown in **Table 4**.

Table 4. Examples of measures and types of qualitative and semi-quantitative assessment results

Assessment Types	Examples of Qualitative or Semi-Quantitative Assessment Results	Examples of Measures
Implementation: Examine the progress of specific controls.	Determine whether identified controls are in place.	The percentage of systems with up-to-date patches (i.e., implementation of a specific control or capability)
Effectiveness: Examine how well controls are working.	Use a color-coded risk matrix to demonstrate the potential risks involved with improperly configured access controls.	A chart that shows the changes of percentage of information security incidents caused by improperly configured access controls over a 5-year period
Efficiency: Examine the timeliness of controls.	Use a 1–5 scale to determine whether the organization is at an acceptable level of responsiveness in case of an information security incident.	Data that compares the mean time of response to information security incidents versus the cost of the incident
Impact: Examine the impact of information security on an organization's mission.	Rank risks on a color-coded scale to evaluate financial impacts to an organization.	Data on the known costs of breaches to industry peers

3.3. Prioritizing Measures

Most organizations have constraints that prevent gathering and analyzing measurement data on every possible measurement data source. For this reason, after implementation measures are in place, organizations prioritize which efficiency, effectiveness, and business impact measures to implement. Prioritization can be driven by a variety of factors, including an organization's risk management strategy, mission and business objectives, information from risk assessments, policies, and legal, regulatory, or other requirements.

3.3.1. Likelihood and Impact Modeling

Likelihood and impact modeling are meant to work in tandem as part of a larger risk assessment process.[11] Simply knowing either the likelihood or the potential impact of an event is not enough information to determine the importance of a potential measure to an organization.

Identifying existing data for use in likelihood and impact modeling typically involves working with stakeholders from across the organizational structure. When possible, data from existing risk assessments can be utilized to reduce redundancy and enable decision-making (e.g., using existing modeled data to help decide what measurements to prioritize). Organizations may also have in-house knowledge from audits, interviews, surveys, or studies that can provide useful data points. In addition to existing internal data, external data on likelihood and impact can be useful as well. Published annual reports can provide information on threat landscapes and the financial impacts of information security incidents that can be used to create models. A wide range of event likelihood models can be used to assess the likelihood of adverse events when determining which systems and controls to measure.

Organizations can also compare impact models with event probability models (e.g., expected loss and statistical analysis of historical market trends) to determine their measurement priorities. Controls or systems with higher likelihoods of incident or higher potential impacts if affected could then be prioritized when organizations decide how to allocate measurement resources. Where possible, leverage existing event likelihood and impact models (e.g., risk registers[12]) to avoid a duplication of efforts. More information on quantitative likelihood and impact models can be found in Appendix C.

In addition to using historical information for likelihood and impact modeling, current trends may provide useful datapoints when prioritizing and selecting measures. Staying up to date on current threats allows for more effective continuous measurement and assessment. At the same time, organizations consider recency bias[13] about current events when determining courses of action and resource allocation. Outliers and unexampled events may occur over time. An organization can prepare for these issues using horizon scanning, stress tests, and system resilience.[14]

3.3.2. Weighing Scale

Information gained from modeling likelihood and impact can be combined with knowledge about organizational goals and existing controls to create a customized weighing scale to prioritize potential measures. Using a weighing scale with set parameters ensures consistency

[11] More information on risk assessments can be found in SP 800-30, *Guide for Conducting Risk Assessments*.
[12] More information on risk registers can be found in [4].
[13] Recency bias is the tendency to favor recent events or experiences over historical ones.
[14] More information on cyber resiliency can be found in SP 800-160 Volume 2.

when prioritizing and selecting measures, even those that are unrelated to information security. Measures that are ultimately selected are useful for:

- Identifying causes of unsatisfactory performance
- Pinpointing areas for improvement
- Facilitating consistent policy implementation
- Redefining goals and objectives
- Modifying security policies

3.4. Evaluating Methods for Supporting Continuous Improvement

After an organization selects its measures, the collected data is evaluated. Evaluation may look different depending on the types of measures being analyzed. Quantitative data analysis methods, like those in Sec. 2.3, can be used to evaluate measures.

For implementation measures, evaluation may be as simple as comparing the percentage of controls implemented with the goal percentage of implementation. Effectiveness, efficiency, and impact measures will likely be more complicated to evaluate. Both effectiveness and efficiency measures often begin by establishing average data output and evaluating acceptable ranges against output going forward. For example, an organization may want to know if the volume of data being transferred on the network has an anomaly. To monitor for changes, the average volume of data transferred is established. An organization may also set an acceptable range based on a standard deviation from this average. This may mean looking for outliers in the data or monitoring for changes over time. Evaluating impact measures will likely include outcomes outside of information security, such as financial outcomes or even public perception.

Various indicators and inputs can be useful to track the effectiveness and efficiency of an information security program by monitoring performance and security over time, such as:

- **False positive rate:** The proportion of positive reports that were incorrectly identified
- **Key performance indicators:** A measure of progress toward intended results
- **Key risk indicators:** A metric used to measure risk
- **Leading indicators:** A predictive metric that tracks events or behaviors that precede incidents
- **Lagging indicator:** A metric that tracks the outcome of events or trends
- **Mean time to detect:** A metric that tracks the average amount of time that a problem exists before it is found
- **Mean time to recovery:** A metric that tracks the average amount of time it takes to recover from a product or system failure

Access to average outputs, acceptable ranges, and long-term data makes effectiveness and efficiency measures more accurate and beneficial by enabling organizations to track changes over time. Even if processes are not yet consistent, average outputs and acceptable ranges help organizations set metrics. Some metrics are directly related to established averages, while others are set by other sources, and established ranges may not have any effect on organizational goals. While inconsistent processes will not provide meaningful data, measurements may still be used to establish average outputs and acceptable ranges for future analysis. Data analysis for finding average outputs and acceptable ranges will typically include historical data and a forecast of what that trend would continue to look like in the future if all variables stay the same.

It is important to remember that some measures have the potential to give misleading information. Inputs such as phishing test success rates or the number of knows vulnerabilities depend heavily on the quality of work behind them. A poorly designed phishing test might show a better success rate while giving less information about the preparedness of the workforce to recognize a well-designed phishing email. This does not mean that organizations need to avoid these measures altogether, but numbers alone may not always show the whole story.

References

[1] National Institute of Standards and Technology (2018) Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Cybersecurity White Paper (CSWP) NIST CSWP 6. https://doi.org/10.6028/NIST.CSWP.6

[2] National Institute of Standards and Technology (2006) Minimum Security Requirements for Federal Information and Information Systems. (U.S. Department of Commerce, Washington, DC), Federal Information Processing Standards Publication (FIPS) 200. https://doi.org/10.6028/NIST.FIPS.200

[3] Bowen P, Kissel RL (2007) Program Review for Information Security Management Assistance (PRISMA). (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) 7358. https://doi.org/10.6028/NIST.IR.7358

[4] Stine KM, Quinn SD, Witte GA, Gardner RK (2020) Integrating Cybersecurity and Enterprise Risk Management (ERM). (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) 8286. https://doi.org/10.6028/NIST.IR.8286

[5] Taylor BN (2011) The current SI seen from the perspective of the proposed new SI. *Journal of Research of the National Institute of Standards and Technology* 116(6):797. https://doi.org/10.6028/jres.116.022

[6] Software Quality Group (2021) Metrics and Measures. (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://www.nist.gov/itl/ssd/software-quality-group/metrics-and-measures

[7] Thomas D (2019). Monte Carlo Tool. (National Institute of Standards and Technology, Gaithersburg, MD). Available at https://www.nist.gov/services-resources/software/monte-carlo-tool

[8] ASTM International (2018) *ASTM C1012/C1012M-18a – Standard Test Method for Length Change of Hydraulic-Cement Mortars Exposed to a Sulfate Solution* (ASTM International, West Conshohocken, PA). https://doi.org/10.1520/C1012_C1012M-18A

[9] Joint Task Force Transformation Initiative (2012) Guide for Conducting Risk Assessments. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-30, Rev. 1. https://doi.org/10.6028/NIST.SP.800-30r1

[10] Joint Task Force (2018) Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-37, Rev. 2. https://doi.org/10.6028/NIST.SP.800-37r2

[11] Chew E, Swanson MA, Stine KM, Bartol N, Brown A, Robinson W (2008) Performance Measurement Guide for Information Security. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-55, Rev. 1. https://doi.org/10.6028/NIST.SP.800-55r1

[12] Grance T, Nolan T, Burke K, Dudley R, White G, Good T (2006) Guide to Test, Training, and Exercise Programs for IT Plans and Capabilities. (National Institute of Standards and

Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-84. https://doi.org/10.6028/NIST.SP.800-84

[13] Kent K, Souppaya M (2006) Guide to Security Log Management. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-92. https://doi.org/10.6028/NIST.SP.800-92

[14] Dempsey KL, Pillitteri VY, Baer C, Niemeyer R, Rudman R, Urban S (2020) Assessing Information Security Continuous Monitoring (ISCM) Programs: Developing an ISCM Program Assessment. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-137A. https://doi.org/10.6028/NIST.SP.800-137A

[15] Barker E, Smid M, Branstad D (2015) A Profile for U.S. Federal Cryptographic Key Management Systems. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-152. https://doi.org/10.6028/NIST.SP.800-152

[16] Ross, R, Winstead M, McEvilley, M (2022) Engineering Trustworthy Secure Systems. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-160v1r1. https://doi.org/10.6028/NIST.SP.800-160v1r1

[17] Boyens JM, Smith AM, Bartol N, Winkler K, Holbrook A, Fallon M (2022) Cybersecurity Supply Chain Risk Management Practices for Systems and Organizations. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-161r1. https://doi.org/10.6028/NIST.SP.800-161r1

[18] Heckert N, Filliben J, Croarkin C, Hembree B, Guthrie W, Tobias P, Prinz J (2012) NIST/SEMATECH e-Handbook of Statistical Methods. (National Institute of Standards and Technology, Gaithersburg, MD). https://doi.org/10.18434/M32189

[19] Stevens SS (1946) On the Scales of Measurement. *Science* 103(2684):677-680. http://www.jstor.org/stable/1671815

[20] Turan MS, Barker E, Kelsey J, McKay KA, Baish ML, Boyle M (2018) Recommendation for the Entropy Sources Used for Random Bit Generation. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication (SP) 800-90B. https://doi.org/10.6028/NIST.SP.800-90B

[21] Stolfo S, Bellovin S, Evans D (2011) Measuring Security. *IEEE Security & Privacy* 9(3):60-65. https://doi.org/10.1109/MSP.2011.56

Appendix A. Glossary

assessment
The action of evaluating, estimating, or judging against defined criteria. Different types of assessment (i.e., qualitative, quantitative, and semi-quantitative) are used to assess risk. Some types of assessment yield results.

assessment results
The output or outcome of an assessment.

Bayesian methodology
Statistical approach to data analysis based on Bayes' theorem where uncertainty is quantified by combining existing information with new information to create forecast models. [18, adapted]

classical data analysis
A data analysis technique where data collection is followed by the imposition of a model and the analysis, estimation, and testing that follow are focused on the parameters of that model. [18, adapted]

data validation
The process of determining that data or a process for collecting data is acceptable according to a predefined set of tests and the results of those tests. [15]

experimentation
A systematic approach to the process of testing new ideas, methods, or activities that applies principles and techniques at the data collection stage to ensure the generation of valid, defensible, and supportable conclusions.

exploratory data analysis
A data analysis technique where data collection is immediately followed by analysis with the goal of inferring what model would be appropriate. [18, adapted]

false positive
An erroneous acceptance of the hypothesis that a statistically significant event has been observed. [20]

imputation
The replacement of unknown, unmeasured, or missing data with a particular value. The simplest form of imputation is to replace all missing values with the average of that variable. [18, adapted]

information security
The protection of information and systems from unauthorized access, use, disclosure, disruption, modification, or destruction to provide confidentiality, integrity, and availability. [2]

interval scale
From the Stevens Scale of Measurement, a quantitative measurement scale using variables with equal values and no true zero, such as time and temperature. [19, adapted]

key performance indicator
A metric of progress toward intended results.

key risk indicator
A metric used to measure risk.

lagging indicator
A metric that tracks the outcome of events or trends.

leading indicator
A predictive metric that tracks events or behaviors that precede incidents.

machine learning
The development and use of computer systems that adapt and learn from data with the goal of improving accuracy.

mean
The sum of the data points divided by the number of data points. Commonly referred to as the average. [18, adapted]

mean time to detect
A metric that tracks the average amount of time that a problem exists before it is found.

mean time to recovery
A metric that tracks the average amount of time that it takes to recover from a product or system failure.

measurement
The process of obtaining quantitative values using quantitative methods.

measures
Quantifiable and objective values that result from measurement.

median
The value of the point that has half the data smaller than that point and half the data larger than that point. [18]

metrics
Measures and assessment results designed to track progress, facilitate decision-making, and improve performance with respect to a set target.

mode
The value of the random sample that occurs with the greatest frequency. This value is not necessarily unique. [18]

Monte Carlo analysis
A probabilistic sensitivity analysis used to account for uncertainty. [7]

nominal scale
From the Stevens Scale of Measurement, a scale that labels named variables into classifications. [19, adapted]

normalization
The conversion of information into consistent representations and categorization. [4]

observational data
Data captured through the observation of an activity or behavior without the direct involvement of the subject.

ordinal scale
From the Stevens Scale of Measurement, a scale that orders and ranks data without establishing a degree of variation between ranks. [19, adapted]

outliers
An observation that lies an abnormal distance from other values in a random sample from a population. [18]

qualitative assessment
The use of a set of methods, principles, or rules for assessing risk based on non-numerical categories or levels. [6]

quantitative assessment
The use of a set of methods, principles, or rules for assessing risk based on numbers where the meanings and proportionality of values are maintained inside and outside of the context of the assessment. [6]

random sampling
A method of sampling where each sample has an equal chance of selection in hopes of gathering an unbiased representation. [18, adapted]

ratio scale
From the Stevens Scale of Measurement, a quantitative measurement scale with a true zero using variables that can be compared to find differences or intervals. [19]

regression
A statistical technique used to predict the value of a variable based on the relationship between explanatory variables.

sampling
The process of taking samples of something for the purpose of analysis.

semi-quantitative assessment
The use of a set of methods, principles, or rules for assessing risk based on bins, scales, or representative numbers whose values and meanings are not maintained in other contexts. [9]

stratified sampling
The process of segmenting a population across levels of some factors to minimize variability within those segments. [18]

systematic stratified sampling
A method of sampling where samples are taken at a regular interval. [18, adapted]

time series analysis
The analysis of an ordered sequence of values of a variable at equally spaced time intervals. [18, adapted]

transformation
The conversion of one state or format into another state or format.

Appendix B. Data Analysis Dictionary

The following information is found in the NIST Engineering Statistics Handbook.

B.1. Bayesian Methodology

Bayesian Methodology consists of formally combining the prior distribution on the parameters and the collected data to jointly make inferences and/or test assumptions about the model of parameters.

- Bayes Formula

$$\mathrm{P(A|B)} = \frac{P(A,B)}{P(B)} = \frac{P(A) \times P(B|A)}{P(B)}$$

- Law of Probability

$$\mathrm{P(B)} = \sum_{i=1}^{n} P(P|A_i)P(A_i)$$

B.2. Classical Data Analysis

Classical data analysis is when data collection is followed by a model, and the subsequent analysis, estimation, and testing focus on the parameters of that model. Classical data analysis includes deterministic and probabilistic models, such as regression and ANOVA. Some of the more common relevant classical quantitative models include:

Location

- Measures of Location (mean, median, and mode)
- Confidence Limits for Mean and One Sample t-Test
- Two Sample t-Test for Equal Means
- One Factor Analysis of Variance
- Multi-Factor Analysis of Variance

Scale (or variability or spread)

- Measures of Scale
- Bartlett's Test
- Chi-Square Test
- F-Test
- Levene Test

Skewness and Kurtosis

- Measures of Skewness and Kurtosis

Randomness

- Autocorrelation
- Runs Test

Distributional Measures

- Anderson-Darling Test
- Chi-Square Goodness of Fit Test
- Kolmogorov-Smirnov Test

Outliers

- Detection of Outliers
- Grubbs Test
- Tietjen-Moore Test
- Generalized Extreme Deviate Test

2-Level Factorial Designs

- Yates Algorithm

B.3. Exploratory Data Analysis

Exploratory data analysis emphasizes graphical techniques and inferring different analytic models in order to determine what model would be appropriate. Some common exploratory data analysis graphical techniques include:

Univariate

$$y = c + e$$

- Run Sequence Plot
- Lag Plot
- Histogram
- Normal Probability Plot
- 4-Plot
- PPCC Plot
- Weibull Plot
- Probability Plot

- Box-Cox Linearity Plot
- Bootstrap Plot

Time Series

$$y = f(t) + e$$

- Run Sequence Plot
- Spectral Plot
- Autocorrelation Plot
- Complex Demodulation Amplitude Plot
- Complex Demodulation Phase Plot
- Decomposition

1 Factor

$$y = f(x) + e$$

- Scatter Plot
- Box Plot
- Bihistogram
- Quantile Plot
- Mean Plot
- Standard Deviation Plot

Multi-Factor/Comparative

$$y = f(xp, x1, x2, \dots, xk) + e$$

- Block Plot

Multi-Factor/Screen

$$y = f(x1, x2, x3, \dots, xk) + e$$

- DOE Scatter Plot
- DOE Mean Plot
- DOE Standard Deviation Plot
- Contour Plot

Appendix C. Modeling Impact and Likelihood

This appendix is intended to provide a high-level overview of complex statistical concepts. The successful application of these concepts will require further training and understanding on the part of practitioners.

C.1. Bayesian Methodology

Bayes' formula expresses the conditional probability of event A given event B written as P(A/B). It can be calculated using Bayes' Rule:

$$\mathrm{P(A|B)} = \frac{P(A,B)}{P(B)} = \frac{P(A) \times P(B|A)}{P(B)}$$

Bayesian methodology is applied when there is previous knowledge of the conditions associated with an event. It can provide conditional probability estimates quickly and without using significant resources. Because Bayesian methodology relies on prior information, it is important to note that the use of either inaccurate or a different selection of prior information may lead to results that do not provide significant insight.

C.2. Monte Carlo Methodology

The Monte Carlo method is a multiple probability simulation used to predict possible outcomes of an uncertain event. The Monte Carlo method uses randomly generated outcomes within a set range, and the frequencies of different outcomes generated form a normal distribution.

The Monte Carlo method allows for repeated modeling and can be performed using spreadsheet editors or programing languages for statistical computing. When using the Monte Carlo method, it is important to note that these simulations show an estimated probability and not an inevitable outcome.

C.3. Time Series Analysis

Time series analysis shows the level, trend, seasonality, or noise within a series of data points in a time series. Time series data is often found when monitoring a process over a period. Time series analysis considers the potential for an internal structure, such as trends or seasonal variations to data.

Time series regression models are primarily used for forecasting. Time series decomposition exhibits patterns within time series data and can be useful when setting the expected range or use of processes or systems.

C.4. Value at Risk

Value at risk (VaR) is a statistical analysis technique that builds a model that measures the risk of loss, primarily using a probability density function. The three key elements of building a VaR model are a fixed time period, a specific level of loss in value, and a confidence interval.

Calculating VaR can be helpful when making decisions about investments and resources. Like all predictive models, VaR relies heavily on the quality of inputs and cannot effectively estimate all scenarios.

Appendix D. Change Log

[Upon final publication, a change log will be included that describes differences from the superseded version of this publication: NIST SP 800-55r1 (2008).]

In <date of final publication> the following changes were made to the report:

- …