

How to Complete the Scope of Data Section of the PMDF

Scope (Range) of Data	?	?	?	?	?	?
-----------------------	---	---	---	---	---	---

The **Scope of Data** section aligns to the V&V criterion of **Complete**

Why the “Scope of Data” important ...



To meet the criterion of Complete, a measure must have a clearly defined scope. The scope tells you the range of data points that are included in the calculation of the measure result. In a more practical sense, scope is about representation – *a defined scope tells you what the reported result represents.*

For example, say you are about to get on a new Rolling Serpent roller coaster. It looks a little scary, so you ask the operator whether it is safe. He reports that 100% of roller coasters consistently pass safety inspections. While initially reassuring, you start to wonder whether that perfect safety inspection score applies to the *new* Rolling Serpent roller coasters. You might also wonder whether the safety inspections covered everything that is likely to go wrong with Rolling Serpent roller coasters especially since the Rolling Serpent is the only type of roller coaster to do an upside-down triple loop.

These are all scope questions. They are important because they tell what the 100% result represents, and more importantly, whether you can use the result to draw conclusions about the safety of the Rolling Serpent roller coaster. A clear scope helps to bound the interpretation of the results.

How to complete the Scope of Data section in the PMDF ...

There are three elements of scope that should be discussed in the Scope of Data section of the PMDF:

- ✓ the **unit of analysis**,
- ✓ the **attribute**, and
- ✓ the **population**.

Note that Scope of Data section must meet the 1,000-character limit set by DHS. Consider the following performance measure used by the XYZ program:

Percent of hazardous waste storage facilities that meet legislated safety standards.

Below is the Scope of Data section for this measure.

Scope (Range) of Data	The unit of analysis is a single hazardous waste storage facility, including all storage buildings, for a given company at a given location. Two companies with facilities at the same location or a single company's facilities at a different location are counted separately. The population includes all domestic facilities inspected within the reporting period, except those with a 401 exemption. A risk-based stratified sample of facilities is selected to be inspected each year with high-risk facilities being inspected each year, medium risk every two years, and low risk every three years. The attribute is whether a given facility passes the safety inspection. Facilities are assessed against 5 critical and 15 noncritical safety standards based on CFR Title 000 Part 000. Each standard is scored as pass or fail. A facility is considered to have passed the inspection if none of the critical standards receive a fail rating and less than five noncritical standards receive a fail rating.
-----------------------	---

The three elements of scope are defined below along with guidance for what to write in the PMDF. As shown in the example above, these elements can usually be **described in three or four sentences**. The guidance below focuses on the content of these sentences.

The **Unit of Analysis** is the individual “thing” – an object, circumstance, event, action, relationship, process, result, etc. – being observed, scored, and counted in order to produce the reported result. The unit of analysis is a single member of the population. *The unit of analysis is important because it identifies the thing on which data is actually being collected and about which conclusions may be drawn from the results.*



Guidance:

- Describe what the unit of analysis is and what it encompasses.
- Provide guidelines for counting a unit of analysis (e.g., *facilities at different sites are counted separately*) and/or how to tell where one unit of analysis ends and another begins (e.g., *each import has its own import entry line*).

The **Population** is the total set of units of analysis that a particular result represents. Unless some form of sampling was used, the population tells you all the units of analysis that were assessed to calculate the measure result. When sampling is used, the population indicates the domain of units of analysis that the sample is assumed to represent. *It is important because it tells you the extent to which you can generalize the conclusions you draw from the reported result.*



Guidance:

- Indicate what the population includes.
- Supplement this information by explaining any limits of the population (e.g., *only facilities above a certain size are included*) and/or units of analysis that are not included in the population (e.g., *facilities with an open inspection at the end of the period are excluded*).
- If a sample is used, explain how the units in the sample were selected.

The **Attribute** is the property or characteristic of the unit of analysis that is being measured – i.e., given a score. The attribute can be any characteristic or property on which the unit of analysis may vary (e.g., height, weight, in compliance, processing time). *It is important because it is the basis for determining whether a particular unit of analysis is counted toward the results.*



Guidance:

- Describe the attribute/characteristic the unit of analysis must possess to be counted in the results. Note that this statement should be consistent with the measure label.
- Explain the range of scores that may be given on the attribute (e.g., *pass/fail, 1 to five, etc.*) and how the scores are assigned to the units of analysis (e.g., *via a safety inspection*).
- Be sure to explain what score (e.g., *must be rated as satisfactory or better*) or combination of scores (e.g., *must pass 8 out 10 standards*) is needed for the unit of analysis to be counted toward the results.

Common Pitfalls



- Unit of analysis does not align with the measure as stated – e.g., *unit of analysis is “Y” but measure states “percent of X”*
- Unit of analysis does not match the use of the measure – e.g., *unit of analysis is “A” but conclusions drawn about “B”*
- Some relevant characteristics of “population” are not defined – e.g., *geographic boundaries, types of organizations or facilities*
- Attribute being assessed is not well defined or lacks a threshold – e.g., *what constitutes “monitored” or “in compliance”*

How to Complete the Data Collection Methodology Section of the PMDF

Measure Description	? ? ? ? ? ?
---------------------	-------------

The **Methodology** section in the PMDF speaks to the V&V criterion of **Complete**.

Why the “Data Collection Methodology” section is important ...



The Data Collection Methodology section of the PMDF is important because it provides an overview of the how, when, where, and who of the procedures and steps by which the data for a measure is collected. Systematic data collection is a fundamental requirement for measurement and is only possible with a well-defined, standardized methodology. Because data collection involves concrete actions and operations carried out by individuals using a combination of procedures, hardware, and IT systems, it brings an element of pragmatism to the measurement process. Data collection is where the theory of measurement meets practical reality. The Data Collection Methodology section of the PMDF documents the viability of the measure.

The Data Collection Methodology section is also important as a crosscheck of whether the conceptual elements of scope align with the concrete actions and operations of data collection. For example, the scope specifies the unit of analysis and attribute of the measure. Does the data collection methodology allow for observation of the unit of analysis and assessment of the attribute? The scope specifies the population. Does the data collection methodology allow for data to be collected across the entire population? The data collection methodology also presents the formula for calculating the results. Does the formula reflect the correct elements of the scope and population to produce the result? These questions help to ensure alignment between the conceptual and practical elements of a measure. Thus, while scope is important, a measure is not “complete” until it has a well-defined and standardized methodology that allows for the collection of data that is aligned with the scope.

How to complete the Methodology section in the PMDF ...



The Data Collection Methodology section should describe the data collection process for the performance measure. Consider the following performance measure used by the XYZ program:

Percent of hazardous waste storage facilities that meet legislated safety standards.

Below is the Data Collection Methodology section for this measure.

Data Collection Methodology	During annual safety inspections, Certified Safety Experts (CSEs) observe and assess whether hazardous waste storage facilities are compliant with safety standards. CSEs record their assessment of each standard, along with any comments, in real time on the Safety Standard Checklist. Within 24 hours, CSEs must enter their ratings into the Facility Information Database (FID). At the end of the quarter, the CO Performance Measure Team (PMT) retrieves the inspection data for all inspected facilities across all field offices for a given period from FID using standardized commands that download the data into a CSV file. The PMT then imports the CSV file into the Inspection Tracker spreadsheet and uses spreadsheet functionalities to analyze the data and calculate the result. The calculation is the number of facilities passing the inspection divided by the number inspected in the period.
-----------------------------	--

In general, it is useful to think in terms of three data collection steps that are common to all performance measures: 1. Observation and Assessment, 2. Data Entry and Retrieval, and 3. Analysis and Calculation. These steps are explained below along with guidance for what to write in the PMDF. Note that your explanation of *how* these steps are carried out should include *when, where, and by whom* as well as any tools or devices used.

1. Observation and Assessment – The Methodology section writeup should describe how, when, where, and by whom the units of analysis are observed and assessed (i.e., given a score) in terms of the attribute. Note that the observation may be fully automated or involve the use of a device (e.g., radar, GPS, satellite, etc.). Nevertheless, it is still important to describe how the assessments of the units of analysis are collected. *The description of the observation and assessment step is important because it helps ensure the methodology is consistent with the scope.*



Guidance:

- Begin your writeup with a sentence that explains how the observation and assessment of the units of analysis takes place (e.g., “Auditors conduct annual onsite review of financial statements”). As needed, indicate how the scores are initially documented (e.g., “Auditors use digital recorders to document their assessments”).

2. Data Entry and Retrieval – The Methodology writeup should continue with an explanation of how the assessment scores are entered into the IT system described in the Data Source section of the PMDF. The writeup should also describe how the data is retrieved for analysis. If a local spreadsheet is used describe how the data is transferred to the spreadsheet. Be sure to describe both manual and automated operations. *The description of the data entry and retrieval is important because it helps to ensure that the raw data is not compromised or changed prior to analysis.*



Guidance:

- Describe how the raw data is entered into the IT system (e.g., “Administrative staff transcribe the recordings onto a local spreadsheet which is then uploaded into the audit database”). Next describe how the data is retrieved (e.g., “an analyst uses a .SQL file to retrieve ...”) and prepared for analysis (e.g., “The SQL creates a CSV file that the analyst manually imports into a local Excel spreadsheet ...”).

3. Analysis and Calculation – The Methodology writeup should describe how the reported result is calculated. As needed, describe any analyses (aggregation, rollups, etc.) executed prior to the calculation of the reported result. Be sure to describe both manual and automated operations. *The description of the calculation is important because it helps to clarify the nature (e.g., average, percentage, count, proportion, frequency, etc.) of the reported result.*



Guidance:

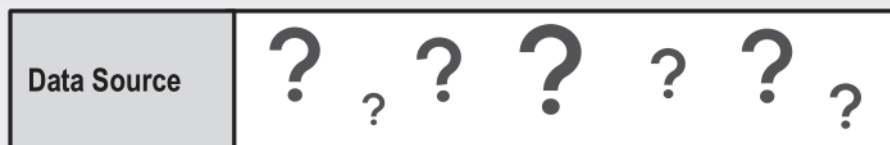
- Describe how the result is calculated (e.g., “HQ analysts use spreadsheet functionalities to calculate the reported percentage”) as well as any analyses performed on the raw data (e.g., “the ratings on each standard are averaged to produce an overall facility score”). Use the last sentence to describe equation used to calculate the results. The equation which must be written in narrative form as the DHS system cannot store a formula.

Common Pitfalls



- The observation and assessment procedures focus on a different unit of analysis than is described in the scope.
- Failing to clearly indicate the data collection operations that are carried out manually.
- The formula is not consistent with the scope.
- Failing to describe when, where, and by whom the data collection steps are carried out.
- Failing to describe analyses or transformations that are carried out on the data before the results are calculated.

How to Complete the Data Source Section of the PMDF



The **Data Source** section in the PMDF speaks to the V&V criterion of **Consistent**.

Why the “Data Source” section is important ...



The Data Source section of the PMDF is important because it provides the reader with an overview of the storage, extraction, and reporting of the measure data and identifies the office(s) responsible for maintaining and reporting the data. Having a defined process and people who oversee it are key elements of the consistency of measurement data. Performance measurement data may be considered consistent if they are stored, extracted, and reported the same way each time. Consistency is more likely to occur when an office is responsible for carrying out those processes. The Data Source section specifies the measurement data, where it is stored, the frequency with which it is updated, queries that are performed, and files to where the data are extracted for analysis and reporting. Describing this information also provides a degree of transparency to the reader.

For example, say a program performance measure generates the result that 80% of domestic airports passed security inspections during the last fiscal year. A question people may ask is “Where did the data for that result come from?” While the Data Collection section of the PMDF will address part of that question, the Data Source section addresses handling of the data once it is collected. The intent is to do that in simple, easy-to-understand terms.

How to complete the Data Source section in the PMDF ...



The Data Source description should provide the reader with key details about the measure’s data. These details primarily concern what the relevant data is, where it resides, how it is updated and extracted, and what offices are involved in maintaining the data system(s) and reporting the data. These are described on the next page. Note that Data Source descriptions must meet the 1,000-character limit set by DHS. Consider the following measure used by the XYZ program:

Percent of hazardous waste storage facilities that meet legislated safety standards.

Below is the Data Source section for this measure.

Data Source	The inspection data for the measure are stored in the Facility Information Database (FID), which contains the size, location, age, and complete inspection record for each facility. Each facility has a unique identification number indicating the type of facility (storage, treatment, or disposal) and whether it handles hazardous waste. The FID is a system of record maintained by the Central Office (CO) IT Group. Monthly, an IT analyst executes a query that complies the field office data and transfers it into the FID. The analyst also exports the data to the Risk Management database owned by the Internal Controls Group. The IC group uses the data to assign the risk ratings that determine a facility’s inspection cycle. At the end of the quarter, the FID data are downloaded into the Inspection Tracker Spreadsheet for the calculation of the measure results. The CO Performance Measure Team (PMT) maintains the tracker and reports the results through the DHS One-number Performance Module.
--------------------	--

The text in the Data Source section should communicate the following details about the performance measure. Each section explains what information should be communicated and provides guidance for what to write in the PMDF.

1. Where the data are stored – The Data Source section should begin with an identification of the systems, spreadsheets, etc. in which the data are stored. This includes the name of the system of record where raw data are stored and any files into which data are transferred or downloaded. *This is important for providing transparency on where data reside and are transferred to during the process of calculating the reported result.*



Guidance:

- Begin with a sentence that identifies the specific system(s) where data are stored (e.g., *Data for this measure are stored in the XYZ system ...*). Spell out systems commonly referred to with acronyms.

2. What data are stored – The Data Source section should also specify what data related to the measure are stored in the system(s). This includes the data on the measure's unit of analysis and attribute. *This is important for understanding what is and is not included in the data systems.*



Guidance:

- Identify data stored in the system (e.g., *"The system contains data on when a facility was inspected, inspection results, location of facilities..."*).

3. Process and timing for transferring data – The Data Source section should continue with a brief description of the processes for transferring data across systems. This includes, as appropriate, describing how and when raw data from the field is compiled into a central system of record. It may also include how and when select data from multiple systems of record are combined into a centralized view or system (e.g., an Enterprise Information System). Finally, it describes how and when data are exported or downloaded into another file for the analysis, calculation, and reporting of measure results. *This is important for understanding the recency of data and how that data is transformed into the reported result.*



Guidance:

- The Data Source should identify when data entered in the field is consolidated into a central database and by whom (e.g., *"On a monthly basis, an IT analyst executes a query that compiles the data from the field offices and transfers it into the XYZ system"*). It should also state any periodic transfer and/or transformation of the data (e.g., *"Each quarter, an analyst exports the data to CFO where it is cross-referenced with financial information and used to analyze and prepare division budgets."*).

4. Office(s) managing systems and reporting results – The Data Source section should specify the office(s) responsible for managing the system(s) and for submitting the reported results to PA&E. *This provides transparency on the roles of those involved in maintaining and reporting the measure data.*



Guidance:

- Identify the office managing the data system(s) (e.g., *"ABC Office manages the XYZ system and downloads data into the ABC spreadsheet"*) and the office reporting the results (*"e.g., DEF Office reports quarterly results"*).

Common Pitfalls



- Not explaining the processes for transferring data to or extracting data from the system of record.
- Not identifying all the data in the system relevant to the measure.
- Not spelling out commonly used acronyms when referring to data systems.
- Not identifying who carries out the various Data Source activities and who owns each database.

How to Complete the Measure Description Section of the PMDF

Measure
Description



The **Measure Description** section in the PMDF speaks to the V&V criterion of **Valid**.

Why the “Measure Description” is important ...



The Measure Description section of the PMDF is important because it provides the reader with an overview of the measure and an explanation of how the measure results are both meaningful and significant. The meaningfulness and significance of the results are key aspects of the V&V criterion of Valid. A performance measure may be considered valid if the information yielded by the measure (i.e., the results) are meaningful for the intended use. The measure description explains how the measure provides an indication of performance and how the results are used to inform internal actions and decisions. In addition, the measure description communicates the significance of the results by explaining how they are related to advancing progress on strategic and programmatic goals. In more practical terms, the Measure Description section tells the reader why it is worth their time to track the results of the measure.

For example, say a program performance measure generates the result that 80% of domestic airports passed security inspections during the last fiscal year. People’s next questions are almost always, “What does that mean and why should I care?” At minimum, people want to know whether that result is good or bad, and what difference does it make to get a good versus a bad result. The purpose of the Measure Description is to provide that information in simple, easy to understand terms.

How to complete the Measure Description section in the PMDF ...



The Measure Description section should provide the reader with key details about the measure. These details primarily concern the interpretation and use of the information produced by the measure and are described on the next page. Note that measure descriptions must meet the 1,000-character limit set by DHS. Consider the following performance measure used by the XYZ program:

Percent of hazardous waste storage facilities that meet legislated safety standards.

Below is the Measure Description section for this measure.

Measure
Description

This measure **assesses effectiveness** at ensuring hazardous waste facilities comply with legislated safety standards. The program **impacts the results** through several activities that impel facilities to comply with safety standards. These include: 1) conducting annual safety inspections according to a risk-based schedule, 2) imposing penalties for noncompliance, and 3) providing coaching and guidance materials at the end of an inspection to assist the facility in reaching compliance. The **results are used** to make changes to the risk-based schedule, assess the efficacy of penalties, and update guidance. This measure **aligns to the agency goal** of maintaining compliance, which **is important** to prevent exposure or release of waste that causes harm to employees and the public, as well as adverse environmental and economic impacts. **External factors** such as economic downturns and turnover of facility safety staff may decrease compliance and lower the results independent of program actions.

The text in the Measure Description section should communicate the following details about the performance measure. Each section explains what information should be communicated and provides guidance for what to write in the PMDF.

1. What the measure assesses – The Measure Description should indicate, in terms that are easily understandable to the public, the aspect of program performance assessed by the measure. *This is important because it helps ensure that conclusions about performance may be drawn from the reported result.*



Guidance:

- Begin your description with a sentence that explains the specific aspect of performance being assessed (e.g., *This measure assesses effectiveness at ...*). Do not use jargon or acronyms.

2. Linkage to program activities – The Measure Description should continue with a brief explanation of how program activities impact the measure results. For the measure results to provide an indication of program performance, the program activity must have a systematic impact on the results. *This is important to ensure that the results provide an indication of performance.*



Guidance:

- The Measure Description should identify programmatic activities that impact the measure results (e.g., *“The program impacts facility compliance by conducting inspections, providing guidance ...”*). If the linkage is not obvious, add a sentence to explain how the activities impact the results (e.g., *“penalties impel compliance and guidance helps facilities understand how to improve”*).

3. Use of the results – The Measure Description should continue with an explanation of how and by whom the measure results are used internally. *This helps to support the validity, interpretability, and usefulness of the measure results.*



Guidance:

- Identify the users of the results (e.g., *senior leadership, field staff, etc.*) and how they use the results (e.g., *inform decisions, make judgements, take actions*).

4. Importance of the measure – The Measure Description should conclude with an explanation of why this activity is important and how it helps demonstrate progress towards strategic and programmatic goals. In other words, why is performing well on the measure is important? *This will help to ensure that the measure helps to communicate the impact of the program.*



Guidance:

- Identify the agency or program goals to which the measure aligns (e.g., *“The measure aligns to ...”*) and why the performance being assessed is important (e.g., *“which prevents harm to the public...”*).

5. External factor affecting the results – As space allows, the Measure Description should identify external factors out of the program’s control and that could adversely affect the measure results.



Guidance:

- List primary external factors that could adversely impact the results (e.g., *“Facility turnover may decrease compliance and lower the results ...”*).

Common Pitfalls



- The description of what the measure assesses is not warranted by the data that is collected – e.g., drawing conclusions about “Y” when the measure assesses “X”.
- The explanation of what the measure assesses is simply a description of the reported result.
- Failing to explain the linkage to program goals and activities.
- Failing to explain the how program activities achieve the outcome when using an outcome measure.