

MLDS CENTER

Maryland Longitudinal
Data System

Better Data • Informed Choices • Improved Results

Data Reporting Standards

April 2015

Version 1.5

Maryland Longitudinal Data System Center

550 West Baltimore Street
Baltimore, MD 21201
410-706-2085

mlds.center@maryland.gov

www.MLDSCenter.org

Maryland Longitudinal Data System Governing Board

William E. "Brit" Kirwan, Ph.D.

Chair

Jennie C. Hunter-Cevera, Ph.D.

Vice Chair

Lillian M. Lowery, Ed.D.

Kelly M. Schulz

David Wilson, Ed.D.

Bernard Sadusky, Ed.D.

Tina Bjarekull

Renee Foose, Ed.D.

Steven Rizzi

Jason Perkins Cohen

Jennifer Strong-Mullinix

Brian Roberts

Ross Goldstein

Executive Director

Lawrence J. Hogan, Jr.

Governor

Data Reporting Standards

Contents

Background	1
Governing Board Recommendations	1
Overview	2
Data Reporting Standards – Outline	3
Data Reporting Standards – Details	3
Step 1: Identify the Question	3
Step 2: Determine Available Data	3
Step 3: Refine the Question and Determine the Population	3
Step 4: Determine Value-Added Benefit	3
Step 5: Sufficiency of Identity Matching	3
Step 6: Robustness of Data Elements	4
Step 7: Verify Results	4
Informing the Public	5
Continuously Improve Data Quality	5

Data Reporting Standards

Background

MLDS receives education and workforce data from the Maryland State Department of Education (MSDE), the Department of Labor, Licensing and Regulation (DLLR), and the Maryland Higher Education Commission (MHEC), who in turn collect that data from a variety of different sources. The quality of the data received by the MLDS Center is dependent on the sources that initially collected the data. Data entry errors or failure to enter critical data fields result in data quality issues that can affect the accuracy of the reporting done by the MLDS Center. Further, because the data that MLDS receives were collected by the agencies from sources for purposes other than longitudinal data reporting, in certain instances it does not include all the data values necessary to support matching across sectors or certain types of reporting.

Governing Board Recommendations

To address issues of data quality and ensure accuracy of information reported by the MLDS Center, the Governing Board made the following recommendation in the *2014 Annual Report to the Governor and General Assembly*.

The accuracy of information reported by the Maryland Longitudinal Data System is of the utmost importance to the Governing Board and the Center. Accuracy is affected by the quality and completeness of the data received, the ability of the Center to match that data across sectors (early childhood, PK-12, higher education, and workforce), and the manner in which the information is presented to the public.

The Governing Board therefore recommends that the Center develop, for Board review and approval, a set of standards and protocols for assessing the accuracy of information reported to the public. The standards should include:

1. An assessment of whether the data relied upon for a report are sufficiently complete to support the information reported;
2. An assessment of whether the information presented can be reconciled against other sources;
3. Criteria for determining whether information based on incomplete data is appropriate to be reported; and
4. Methods for informing the public regarding the information published by the Center.

The Governing Board recommends that these standards and protocols be established prior to the release of public information from the MLDS.

Data Reporting Standards

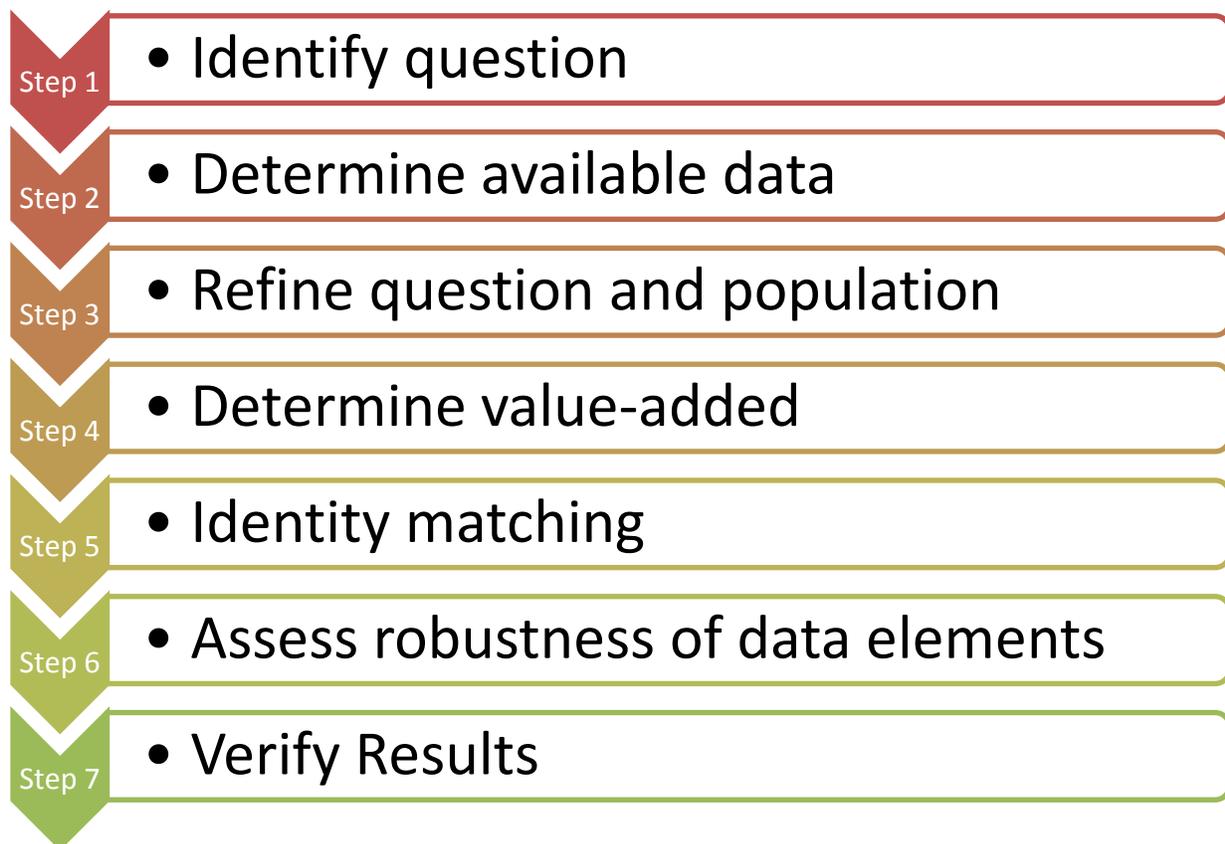
Overview

The following are standards that the MLDS Center will use prior to releasing web content—dashboards with tables, graphs, or other figures—that are based on Center data. The goal of these standards is not to eliminate reporting of information in instances where data are of limited quality. Instead, the goal of these standards is to assess the data limitations and the impact of those limitations on reporting, but still find ways to provide valid, credible, and meaningful information from the data.

The Data Reporting Standards were developed as a step by step process in which each step builds on the next. The Data Reporting Standards are designed so that at each step in the process can be evaluated and refined.

Figure 1, provides a high level description of the steps in the Data Reporting Standards. A detailed description of each of the steps is provided in the next section.

Figure 1: Data Reporting Steps—Simplified



Data Reporting Standards – Details

Step 1: Identify the Question

The Center will identify the initial question based on the *Research Agenda* approved by the Governing Board and in accordance with the purpose of the Maryland Longitudinal Data System according to Maryland Education Article 24-702.

Step 2: Determine Available Data

The Center will determine the data that may be needed to answer the question detailed in Step 1 and will determine the scope of the data available to the Center. Center staff will use the MLDS Data Inventory to determine which data elements are available in the System and for what populations those data elements are available. For the purposes of the System, a data element is any piece of information that may go into an analysis of the data. Data elements may be pieces of information within one point in time (e.g., grades for math and English) or pieces of information over multiple points in time (e.g., absences for each school year).

Step 3: Refine the Question and Determine the Population

The Center will refine the initial question based on the data available in the System as determined in Step 2, specifically define the subset of the population in which data are available to be analyzed.

Step 4: Determine Value-Added Benefit

The Center will determine if the same analysis has been done with the same or similar data elements and subset of the population.

If the analysis has been done, the Center will not continue with the analysis and return to Step 3 in order to revise the question in a way that provides new or additional information.

If the analysis has not been done, continue to Step 5.

Step 5: Sufficiency of Identity Matching

The Center will determine if the source records include the necessary valid data values to support matching person identities within the system.

If ten percent (10%) or more of the data variable necessary to support matching are missing, the Center will conduct a careful examination to discern:

1. What information is missing;
2. Whether there is a pattern to the missing information;
3. What corrective steps may be taken to either obtain the missing information; and
4. What changes to the reporting may be made so that the Center may proceed to use the data and correctly specify the limitations or exceptions thereto.

Data Reporting Standards

Step 6: Robustness of Data Elements

The Center will determine if the source records include valid and complete data on the data elements for the variables, within and across years, needed to provide a complete report.

1. If ten percent (10%) or more of the data values for any element necessary for the analysis are missing or questionable, conduct a careful examination to discern:
 - a. What information is missing;
 - b. Whether there is a pattern to the missing information;
 - c. What corrective steps may be taken to obtain the missing information; and
 - d. What changes to the reporting may be made so that the Center may proceed to use the data and correctly specify the limitations or exceptions thereto.

Step 7: Verify Results

The Center will examine other published sources, especially those produced by partner agencies but also including those produced by educational or research institutions and other government agencies, to determine whether this data or substantially similar data have been published elsewhere.

If the MLDS findings directly contradict findings published by other sources, the Center findings will not be published, unless it is determined that:

1. The Center's findings are judged to be more valid and credible than the other source; or
2. The information can be reported in part through reducing the unit of reporting analysis, eliminating segments of the total population of concern, or by representative sampling.

If the MLDS findings and the published sources differ for methodological reasons, these reasons should be identified and clearly communicated in MLDS products. Some reasons may include:

- Different terminology;
- Different data sources; or
- Different units of inquiry.

If the MLDS findings cannot be compared to other published sources, the Center will determine if the findings are valid and credible enough to be reported based on the following:

1. What is known about the data (Step 2)
2. The addition of new information (Step 4)
3. The sufficiency of matching identities (Step 5)
4. The robustness of the data provided (Step 6)

If differences in the findings cannot be resolved through the options above, the Center will not release the findings.

Data Reporting Standards

Informing the Public

All data dashboards and data snapshots on the Center website will contain a clear description of the relevant steps of these Data Reporting Standards. The Center may provide the following when publishing data publicly when relevant:

- Glossary of Terms
- Calculations of metrics
- Data parameters
- Methodology
- Statistical analyses
- As-of-Date for reported data
- Sources of reported data
- Timeframes of reported data
- Data limitations

Continuously Improve Data Quality

Matching data from multiple sources is essential to the mission of the MLDS Center. Therefore, the MLDS Center will continuously review and revise its matching procedures and the data that are used for matching, in an effort to ensure the highest possible match rate.

For each data source received from a partner agency, the MLDS Center will report the percentage of submitted records with complete and valid data elements in the fields essential to identity matching. These include:

- Social Security Number;
- State Assigned Student Identifier (SASID);
- First and last names;
- Date of birth;
- Race or Ethnicity; and
- Gender.

The Center will review its identity matching processes periodically to determine whether different processes will result in additional matches.

The Center will examine data contributed by its partner agencies and work with partner agencies to assist them in identifying and improving gaps in data quality. Examples of data problems include:

- Inconsistencies in how data are reported;
- Percentages of missing or invalid data; and
- Consistent patterns of invalid or missing data.

Data Reporting Standards

The Center will share findings on data gaps with partner agencies to assist with agency data quality improvement.

Data Reporting Standards

Document History

Date	Document Version	Document Revision Description	Document Author
February 2015	1.0	Creation based on Data Reporting Standards drafts	Laia Tiderman Ross Goldstein
April 2015	1.5	Revisions based on suggestions from the Regional Education Laboratory, Longitudinal Data User Resource Alliance	Ross Goldstein Angela Henneberger Laia Tiderman

Content Contacts:

Ross Goldstein
Ross.Goldstein@maryland.gov
 410-706-2085

Laia Tiderman
Laia.Tiderman@maryland.gov
 410-767-0037

Angela Henneberger
Angela.Henneberger@maryland.gov
 410-706