Weeding Out Errors in Your Data
Presenters

Fort Worth, TX – March 3–4, 2020
Tony Ruggiero, IDEA Data Center

Nashville, TN – Cancelled Due to COVID-19
Tiffany Boyd, IDEA Data Center
Agenda

• Importance of high-quality data
• Defining high-quality data
• Strategies for improving data quality
• High-quality data and business rules relationship
• Importance of business rules
The Growing Importance of High-Quality Data

• Shift from monitoring to Results Driven Accountability
• Support increasing data use and building data culture
• Critical for informed decisionmaking
• Organizational asset
High-Quality Data Definition

• Timely
  - Current, submitted on time, and still relevant

• Accurate
  - Reliable and consistent across time, locations, and methods

• Complete
  - Entire population represented; no missing records
Data Quality Quick Fix

The weed killer approach

• Cleansing the data at the SEA level after the LEA submissions are finalized

• No feedback loop back to those submitting the data

• The same types of errors happen year after year
Addressing Data Quality at the Roots

Getting to the root of the problem by preventing errors before they happen through

- Data system edit checks
- LEA data validation and review
- Training of LEAs on data requirements and data quality review
Risks of Poor Quality Data

• Undermine credibility and breed organizational mistrust
• Make it harder to identify areas for focused effort for improvement
• Inhibit good data-driven decisionmaking
Strategies for Improving Data Quality

- **Prevent**
  - At time of data entry
  - Restrict values

- **Detect**
  - During the data quality review process

- **Repair**
  - Discovered after data are used
  - Find source of error and implement strategies to prevent in the future
Relationship between high-quality data and business rules

- Business rules, or edit checks, prevent errors at the point of data entry
- Using edit checks is not a foolproof way to obtain high-quality data, but edit checks are part of a data cleaning process
- Program staff and those responsible for the data system edit checks must work together
Examples of Business Rules

- Ensure values are within a given range or set of permitted values
- Ensure mandatory data elements are provided
- Determine if data are in the correct format
- Check that values are consistent across data elements
- Perform year-to-year comparisons
Detect Missing Data or Null Values

Ensure mandatory fields are provided such as

- Disability status
- Birthdate or age
- Grade level
- Race/ethnicity
- Gender
- Enrollment date
Prevent Invalid Formats

Enforce proper formatting of data such as

- Date format: *mm/dd/yyyy*
- Decimals such as #.##
Weed Out Invalid Combinations

• Verify educational environment is valid based on student’s age
• Verify the age is between the state-defined range for students coded with a disability category of Developmental Delay
Enforce Proper Procedures

Rules can help LEAs and schools follow proper procedures, such as ensuring that a student cannot have a basis of exit of “return to regular education” if there is not an end date on the child’s individualized education program (IEP) record.
Importance of Business Rules Documentation

• Communication tool for stakeholders, data submitters, and data users
• Support to develop common understanding
• Essential step in the process
• Reference for new and existing staff
Sharing Business Rules Documentation

Sharing documentation by
• Posting on public website
• Providing with data requests
• Posting on internal shared drives

Sharing documentation will inform data users about
• The number and types of validity checks
• The rigor of data quality checks

This will result in the data users having more confidence in the data.
Questions
Contact Us

• Tiffany Boyd, tiffany.boyd@aemcorp.com
• Tony Ruggiero, tony.ruggiero@aemcorp.com
For More Information

• Visit the IDC website
  http://ideadata.org/

• Follow us on Twitter
  https://twitter.com/ideadatacenter

• Follow us on LinkedIn
  http://www.linkedin.com/company/idea-data-center
This presentation was supported by a grant from the U.S. Department of Education, #H373Y190001. However, the contents do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the federal government.

**Project Officers:** Richelle Davis and Rebecca Smith