



Data Quality Action Plan

Common Data Quality Challenges and a
Sample Action Plan

Practice Manager

Chetan Rao

Contributing Consultants

Leonor Keller

Ben McGuire

Julia Linfors

Gwen Perrin

LEGAL CAVEAT

EAB is a division of The Advisory Board Company. The Advisory Board Company has made efforts to verify the accuracy of the information it provides to members. This report relies on data obtained from many sources, however, and The Advisory Board Company cannot guarantee the accuracy of the information provided or any analysis based thereon. In addition, The Advisory Board Company is not in the business of giving legal, medical, accounting, or other professional advice, and its reports should not be construed as professional advice. In particular, members should not rely on any legal commentary in this report as a basis for action, or assume that any tactics described herein would be permitted by applicable law or appropriate for a given member's situation. Members are advised to consult with appropriate professionals concerning legal, medical, tax, or accounting issues, before implementing any of these tactics. Neither The Advisory Board Company nor its officers, directors, trustees, employees, and agents shall be liable for any claims, liabilities, or expenses relating to (a) any errors or omissions in this report, whether caused by The Advisory Board Company or any of its employees or agents, or sources or other third parties, (b) any recommendation or graded ranking by The Advisory Board Company, or (c) failure of member and its employees and agents to abide by the terms set forth herein.

The Advisory Board Company, EAB, and Education Advisory Board are registered trademarks of The Advisory Board Company in the United States and other countries. Members are not permitted to use this trademark, or any other trademark, product name, service name, trade name, and logo of The Advisory Board Company without prior written consent of The Advisory Board Company. All other trademarks, product names, service names, trade names, and logos used within these pages are the property of their respective holders. Use of other company trademarks, product names, service names, trade names, and logos or images of the same does not necessarily constitute (a) an endorsement by such company of The Advisory Board Company and its products and services, or (b) an endorsement of the company or its products or services by The Advisory Board Company. The Advisory Board Company is not affiliated with any such company.

IMPORTANT: Please read the following.

The Advisory Board Company has prepared this report for the exclusive use of its members. Each member acknowledges and agrees that this report and the information contained herein (collectively, the "Report") are confidential and proprietary to The Advisory Board Company. By accepting delivery of this Report, each member agrees to abide by the terms as stated herein, including the following:

1. The Advisory Board Company owns all right, title, and interest in and to this Report. Except as stated herein, no right, license, permission, or interest of any kind in this Report is intended to be given, transferred to, or acquired by a member. Each member is authorized to use this Report only to the extent expressly authorized herein.
2. Each member shall not sell, license, republish, or post online or otherwise this Report, in part or in whole. Each member shall not disseminate or permit the use of, and shall take reasonable precautions to prevent such dissemination or use of, this Report by (a) any of its employees and agents (except as stated below), or (b) any third party.
3. Each member may make this Report available solely to those of its employees and agents who (a) are registered for the workshop or membership program of which this Report is a part, (b) require access to this Report in order to learn from the information described herein, and (c) agree not to disclose this Report to other employees or agents or any third party. Each member shall use, and shall ensure that its employees and agents use, this Report for its internal use only. Each member may make a limited number of copies, solely as adequate for use by its employees and agents in accordance with the terms herein.
4. Each member shall not remove from this Report any confidential markings, copyright notices, and/or other similar indicia herein.
5. Each member is responsible for any breach of its obligations as stated herein by any of its employees or agents.
6. If a member is unwilling to abide by any of the foregoing obligations, then such member shall promptly return this Report and all copies thereof to The Advisory Board Company.

Introducing Your Data Quality Action Plan

The Data Quality Challenge

Across higher education, the value of data in institutional strategy is increasing rapidly as pressures on traditional revenue sources and unsustainable cost growth make efficiencies, alternative revenue streams, and internal reallocation key to success. However, investment and effort to support new analytics must be paired with improvements in the underlying foundational data (e.g., faculty activity, classroom utilization). Much of this information remains uncollected, inconsistent, and impossible to aggregate across units. Even member institutions that have identified these challenges and fixed current data often struggle to build a sustainable infrastructure to allow for ongoing identification and remediation of inaccurate data.

This document defines the most common data challenges identified through EAB's work with members, and their prevalence among the institutions with whom we partner. It then describes how a sample institution ("Alpha University") might approach their most pressing issues, considering the following four principles:

- 1 Identifying Who Needs to Participate in Discussion**

Get the right people in the room to define critical terms and agree on data policies

 - Who sits on standing committees tasked with data governance over key terms?
 - Should anyone else also be included in these conversations?
- 2 Defining Key Terms Consistently**

Establish shared definitions for foundational data terms

 - How do we define critical terms? Do all units define them in the same way?
- 3 Data Quality Process Checklist**

Identify gaps in data creation and maintenance processes

 - Which standard, strong, and advanced processes do we complete?
- 4 Establishing Processes for Sustainable Quality**

Identify protocols and responsibilities for ongoing data quality assurance

 - How often do we review key data elements for consistency and accuracy?
 - Who are the college- and campus-level data stewards?

Common Data Quality Challenges Across Higher Ed

	Data Quality Challenges	Definition	Prevalence in Higher Education
Student Information System	Attributes of Students Not Identified	The institution should collect key details (e.g., race, ethnicity, gender, socioeconomic status) of all students using consistent and shared definitions; in many cases, this data collection is confined to incoming freshman students only	
	Incorrect Faculty Percent Responsibility	Faculty share of a co-taught course should reflect actual percent responsibility for classes taught (and, separately tracked, percent of classes assigned); percent responsibility across multiple faculty for one course section should equal 100%	
	Faculty Not Matched to Courses They Teach	The faculty identifier tied to a specific course section and term must reflect the actual faculty member teaching the course; in many cases, placeholders are used when teaching faculty are not yet assigned, or one ID is used for multiple graduate students	
	Sections Assigned to Wrong Course Types	All sections taught should be tied to consistent definitions that reflect the teaching style of that section; in many cases, course types are set at default options (i.e., lecture) instead of reflecting diverse styles of instruction	
	Enrollment Caps Set to 'Zero' or Inaccurate	All sections should be set to a number greater than zero; in many cases, faculty use zero caps as a shortcut to approve students, which prevents foundational analyses on class size, utilization, and faculty workload	
	Cross-listed Courses Not Connected	When a course is cross-listed and composes multiple sections from departments, the institution should designate those sections clearly; in many cases, cross-listed courses are not linked in any data field and must be manually connected for analysis	
	Student Majors Not Tracked Across Time	Student majors should accurately be tracked when they are changed; in many cases, student major changes may only be input several semesters after the change was made (e.g., to finalize graduation processes)	
	Out-of-Date Academic Organizational Schema	Academic organizational schema can change dramatically over time, so consistent tracking of units across time is crucial to preserving a longitudinal view of institution trends	
	Online/Hybrid Courses Inaccurately Listed	Each college and department should use the same set of terms to define and describe courses with mixed and online modalities; in many cases, these definitions are inconsistent and courses are assigned incorrect course type codes	
HR System	Decentralized Release Time Data	Information on faculty activities in 'release' from instruction should be defined and tracked consistently across the institution; in most cases, these activities are not tracked outside of department- or college-level administration	
	Inaccurate Tenure and Rank Codes	Tenure/rank uniquely group faculty and instructional staff into meaningful categories describing their relationship with the institution; this information is often not tracked, outdated or simply inaccurate	
	Nonstandard Employee Position Titles	Employee position titles should reflect employee pursuits in a standardized and easy to categorize way; open-text fields result in meaningless classifications that hinder analysis	
	No 'Home' Department Designated for Faculty	Employees at the institution, in particular those with instructional responsibilities, should have a 'home' department; home department generally reflects the administrative unit that is responsible for their career at the institution	
	All Employees Not Accounted For	All employees at the institution should be centrally tracked in the HR employee and payroll files; some staff (e.g. graduate assistants and teaching assistants) are never entered into HR files, preventing analysis of teaching roles and workloads	
Finance System	Uncredited Faculty PI Assignments	The institution should always know which individual(s) on campus is responsible for administering new research grants and all grants must clearly identify investigators; in many cases, grants do not clearly identify principal investigators	
	Research Grant Details Not Captured	All research grants tracked in the central grant ledger should include FOAPAL details, research sponsors, research categories, award amounts, and cost recovery; in many cases, these details either do not exist or are not centrally collected	
All Systems	Misaligned Financial & Academic Orgs	Financial and academic organizational structures should map to each other in order to seamlessly track academic unit level costs in the operating ledger	

Applies to few members
 Applies to half of members
 Applies to most members
 Applies to virtually all members

Source: Analysis of 50+ EAB member institutions during data validation and configuration by data analysts; Conversations with local data experts on critical challenges.

Data Quality Improvement Opportunities

Data Quality Improvement Templates

1 Assigning Sections to the Correct Course Types



- Necessary to accurately track faculty instructional workload and associated costs dedicated to different course types
- Necessary to determine student success effects (i.e., course-level completion) of course type variation and instructional methods

2 Cross-walking Financial and Academic Orgs



- True tracking of academic revenues and costs is not possible without a consistent framework for tying decisions from academic stakeholders to their budgetary impact; cost analyses (in particular cost per SCH analyses) to support budgeting and resource allocation decisions are contingent on a strong and accurate crosswalk

3 Establishing Accurate Enrollment Caps



- Necessary to accurately track section and space utilization rates; without an accurate measure of the theoretical maximum capacity of a course, it is impossible to determine trends in course size, identify the optimal room for that course, and whether the course is under-filled or over-utilized.

Assigning Sections to the Correct Course Types

Preparing Stakeholders and Defining Key Terms

Definition: All sections taught should be tied to consistent definitions that reflect the teaching style of that section; in many cases, course types are set at default options (i.e., lecture) instead of reflecting diverse instruction styles.

System: SIS

Importance of This Data

- Necessary to accurately reflect instructional workload and cost variations tied to different course types (e.g., independent study, seminar, lecture)
- Necessary to determine student success implications of different pedagogical styles

1 Identifying Who Needs to Participate in Discussion

Is there a standing campus committee which is engaged in data governance over course types? _____

If not, which individuals on campus should have input into course types discussion (e.g., representative of the faculty senate, representative from the provost's office, IR and IT, registrar)? If a standing committee exists, use this space to suggest other individuals on campus whose input may be valuable. _____

2 Defining Key Terms Consistently

Terms to Manage

Independent Study

– _____

Seminar

– _____

Lecture

– _____

Lab

– _____

Questions

- What is our process to propose and approve available course types which faculty can choose from?
- Does every college and department understand and use these terms in the same way?
- Where can we identify confusion and/or disagreement about key definitions?
- Who (e.g., chair, dean) approves course type assignments?

Assigning Sections to the Correct Course Types

Establishing an Effective and Sustainable Process

3 Data Quality Process Checklist

Those answering "Yes" are engaged in the described practice

		Yes	No
<input type="checkbox"/>	Faculty/department staff choose from a list when assigning course types rather than entering in a free-text field.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	There is a single, consistent set of course type codes that apply to each course taught in the last semester.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	There is a process in place to identify when course size is out of alignment with course type (e.g., a five-person lecture or 200-person seminar).	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	There are established campus- and college-level guidelines governing the percentage of courses that should be taught in each course type.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	The registrar regularly audits course type data to ensure that course types are accurately assigned based on course size.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	When data stewards identify incorrect course type data, they reach out directly to responsible faculty for one-on-one education on correct data entry.	<input type="checkbox"/>	<input type="checkbox"/>

Standard Practice
 Strong Practice
 Advanced Practice

4 Establishing Processes for Sustainable Quality

How often should we review course type code data: _____

Who is responsible for reviewing, correcting, and verifying course types data:

Campus Data Steward: _____

Questions

- What is our process to review course types data?
- How often do we centrally examine and rectify incorrect course types?
- Who is responsible for reviewing and ensuring accuracy of course types?

Cross-walking Financial and Academic Organizations

Preparing Stakeholders and Defining Key Terms

Definition: Financial and academic organizational structures should map to each other in order to seamlessly track academic unit level costs in the operating ledger

System: Finance, SIS, HR

Importance of This Data

- True tracking of academic revenues and costs is not possible without a consistent framework for tying decisions from academic stakeholders to their budgetary impact; cost analyses (in particular cost per SCH analyses) to support budgeting and resource allocation decisions are contingent on a strong and accurate crosswalk

1 Identifying Who Needs to Participate in Discussion

Is there a standing campus committee which is engaged in data governance over academic-financial crosswalks? _____

If not, which individuals on campus should have input into crosswalk discussion (e.g., representative of the faculty senate, representative from the provost’s and chief business officer’s office, IR and IT)? If a standing committee exists, use this space to suggest other individuals on campus whose input may be valuable. _____

2 Defining Key Terms Consistently

Potential Department-Organization Misalignment Areas

Department	Organization Code
College of Education	-
Educational Leadership & Policy Studies	-
School of Clinical Health Professions	-
Environmental Health & Clinical Lab	-
College of Arts & Sciences	-
College of Justice & Safety	-
English & Theatre	-
Art & Design	-
College of Communications	-

Questions

- What is the relationship between academic departments and organizational codes in our general and operating ledgers?
- Does every academic department have a one-to-one relationship with one organizational code?
- What explains those cases when one department has multiple org codes, or one org code has multiple departments?

Cross-walking Financial and Academic Organizations

Establishing Effective and Sustainable Process

3 Data Quality Process Checklist

Those answering "Yes" are engaged in the described practice

		Yes	No
<input type="checkbox"/>	The list of current academic departments in academic systems is accurate to our current courses taught.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	The list of current organizational codes reflects all effective entities on campus which conduct operations on campus.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	Every department is matched one-to-one with a single organizational code.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	When academic and/or financial organizational schema change, there is a process in place to update crosswalks to reflect new budget realities.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	A team monitors for inconsistencies in academic-financial crosswalks at least once per term.	<input type="checkbox"/>	<input type="checkbox"/>

Standard Practice
 Strong Practice
 Advanced Practice

4 Establishing Processes for Sustainable Quality

How often should we review academic-financial crosswalks: _____

Who is responsible for reviewing, correcting, and verifying organizational crosswalks:

Campus Data Steward: _____

Questions

- What is the relationship between academic departments and organizational codes in our general and operating ledgers?
- Does every academic department have a one-to-one relationship with one organizational code?
- What explains those cases when one department has multiple org codes, or one org code has multiple departments?

Establishing Accurate Enrollment Caps

Preparing Stakeholders and Defining Key Terms

Definition: All sections should be set to a maximum set by pedagogical and space constraints; in some cases, faculty use zero caps as a shortcut to approve students, which prevents foundational analyses on class size, utilization, and faculty workload. In many areas, enrollment caps are simply inaccurate or not tracked.

System: SIS

Importance of This Data

- Necessary to accurately track section and space utilization rates; without an accurate measure of the theoretical maximum capacity of a course, it is impossible to determine trends in course size, identify the optimal room for that course, and whether the course is under-filled or over-utilized.

1 Identifying Who Needs to Participate in Discussion

Is there a standing campus committee which is engaged in data governance over enrollment maximums? _____

If not, which individuals on campus should have input into enrollment caps (e.g., representative of the faculty senate, representative from the provost's office, IR and IT, registrar)? If a standing committee exists, use this space to suggest other individuals on campus whose input may be valuable. _____

2 Defining Key Terms Consistently

Terms to Manage

- Minimum Enrollment Cap for Any Course: _____
- Minimum Enrollment Cap for Seminar Course: _____
- Maximum Enrollment Cap for Seminar Course: _____
- Minimum Enrollment Cap for Lecture Course: _____
- Maximum Enrollment Cap for Lecture Course: _____
- Percentage of Caps Currently Set to 0: _____
- Percentage of Caps Currently Set to 1: _____
- Percentage of Caps Currently 2-9: _____
- Percentage of Caps Currently 10-19: _____
- Percentage of Caps Currently 20-49: _____
- Percentage of Caps Currently 50+: _____

Questions

- What process must faculty complete to designate the size of a course? Is this process consistent across colleges?
- What is the role of pedagogy versus space versus historical trend in determining enrollment cap?
- Do department chairs and/or deans audit enrollment caps for space or cost reasons?
- Who approves final enrollment caps on a term-by-term basis?

Establishing Accurate Enrollment Caps

Establishing Effective and Sustainable Process

3 Data Quality Process Checklist

Those answering "Yes" are engaged in the described practice

		Yes	No
<input type="checkbox"/>	Faculty who wish to approve all students enrolling in a course can access a simple process to do so through the central registrar's office.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	The number of enrollment caps that are set to zero each term is tracked, as is how that number has changed over time.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	Colleges and departments that have zero enrollment caps assigned, and which course types are commonly set to zero, are regularly audited.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	A campus-wide policy exists restricting any enrollment maximums set to zero supported by the provost and faculty senate.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	A campus-wide agreement exists around enrollment cap definitions and utilization.	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	We have made it impossible to set enrollment caps to zero in the course registration system.	<input type="checkbox"/>	<input type="checkbox"/>

Standard Practice
 Strong Practice
 Advanced Practice

4 Establishing Processes for Sustainable Quality

How often should we review enrollment cap data: _____

Who is responsible for reviewing, correcting, and verifying enrollment cap data:

Campus Data Steward: _____

Questions

- What is our process to review zero/inaccurate (i.e., dramatically under- or over-filled) enrollment caps?
- How often do we identify zero caps and follow up with the departments/faculty members responsible?
- Who is responsible for reviewing enrollment caps and ensuring that they are set accurately?

Guidelines for Improving Data Entry Processes

Checklist for Cleaner Data

- _____ **1 Identify data elements that have recurring quality issues**

Methods to identify the “usual suspects” for poor-quality data include:

 - Comparing results when different departments pull similar reports
 - Meeting with IR staff to discuss data discrepancies they’ve witnessed
 - Asking department heads which data they trust least because of potential quality issues

- _____ **2 Identify campus members who are knowledgeable about the related data entry processes and how the data is used for decision making**

The following campus members may possess useful information on these issues:

 - Data entry staff and their managers
 - Campus members who use the data to inform decisions

- _____ **3 Determine current practice for data entry in different departments across the institution and identify the user requirements for data use**

- _____ **4 Fix system problems that lead to poor-quality data entry**

Potential workarounds or other solutions may include:

 - Converting open fields to drop-down menus
 - Limiting the range of acceptable numerical entries
 - Providing temporary solutions for unknown values, to be fixed later

- _____ **5 Assign one campus member to develop a standardized process for data input and maintain responsibility for the process moving forward**

Potential candidates include:

 - Business intelligence analyst
 - Director of the department that performs the most relevant data entry

Decisions to make include:

 - Who owns responsibility for relevant data entry
 - What valid entries exist for each relevant data element (format and potential values)
 - What fields are required versus optional
 - How staff should manage required but unknown fields

- _____ **6 Document new processes, distribute agreed-upon standards to all relevant constituents, and store instructions in a centrally accessible location**

Ensure that the following campus members receive a copy:

 - All staff responsible for related data entry
 - Data custodians or data stewards of the relevant data

- _____ **7 Create accountability mechanisms to ensure high data quality over time**

Potential accountability mechanisms include:

 - Automated data quality checks and communication of results to data entry staff
 - Department-level data quality scorecards



EAB

2445 M Street NW, Washington DC 20037
P 202.266.6400 | F 202.266.5700 | eab.com