

# Archived Information

---

## Education Department's Draft Data Quality Standards

---

### **Standard 1: Validity**

1. Do the objective, performance indicator, and data all describe the **same** phenomena of interest? Do they all align?
2. Is a realistic plan in place to improve data collection on all performance measures (especially to resolve any mismatches addressed in question #1)?
3. Do the indicators cover aspects of the program that are useful and important for policy decision making?
4. Is way in which the data are being used to measure performance on the indicator appropriate given what is known about how it was collected?

### **Standard 2: Accurate Definitions**

1. Have clear written definitions of key terms (including exclusions/inclusions) been communicated to data providers?
2. Have definitions and time periods been followed or discrepancies clearly documented?
3. Do reporting forms provide spaces for data providers to report deviations from definitions and uses of estimation at the time they provide the data?
4. Have definitions been communicated in sufficient time for data providers to prepare their system to properly implement them?
5. Has feedback been invited from data providers about data collection issues and possible problems?
6. Have respondents been involved in setting definitions for key terms?

### **Standard 3: Accurate Counts**

1. Have entities for which counts have changed more than 10% since the previous report been double-checked?
2. Have estimates been used for no more than 10% of the phenomena counted, and are estimates clearly differentiated from actual counts?
3. Have counts been tallied at least twice and totals agree?
4. Are independent under and overcount checks in place for counts associated with major program funding?
5. Has sample been drawn from the most up-to-date population lists available?
6. Have weights been properly applied?
7. Are data reported with weights properly applied?

### **Standard 4: Editing**

1. Have non-responses been followed up?
2. Have large changes or unusual findings been discussed with the primary data providers to see if they might be due to editing errors?
3. Have data errors been traced back to their original source and mistakes corrected?
4. Have the data been "eyeballed" to see if it is reasonable given what is known about earlier years and other respondents?

5. Have the data, after they were entered, been systematically reviewed by a different person who is familiar with the data?
6. Has an electronic editing program been used to clean the data?

### **Standard 5: Calculation**

1. Have procedures for dealing with missing data been correctly applied?
2. Have the “+ or –“ confidence intervals been reported for sample data?
3. Have systems been put in place to double-check all data?
4. For sample data, has the data analysis plan been reviewed by outside experts to ensure that appropriate formulae and procedures are applied?

### **Standard 6: Timeliness**

1. Are data from a time before the policy period of interest?
2. Is a regularized schedule of data collections in place to meet policy information needs?
3. Are data reported as soon as possible after collection?
4. Is the year of the data collection clearly identified in the report?
5. Are the data entered and processed in electronic machine readable form?
6. Are resources being invested in creating a modern automated electronic data system?
7. Are review processes designed to ensure that findings are made public in a timely fashion?
8. Are respondents involved in setting time schedules?
9. Are time schedules for providing data enforced with clear and frequent reminders?

### **Standard 7: Reporting**

1. Are data quality problems reported together with the findings?
2. Are data quality problems clearly described in reports?
3. If there have been significant changes in program definitions that might break trend lines, have they been noted?
4. Is the data collection method and sample size mentioned at least briefly when findings are presented?
5. Is the year that the data were collected clearly stated in the report?
6. Are data collection, cleaning, and analysis procedures documented in writing?
7. Is each step in the data collection process required to report deviations and problems in data quality?
8. Are good graphics techniques used (e.g., axes at 0) ?
9. Have the types of exclusions and amount of non-response been clearly described?
10. Has the report of findings been edited by someone with expertise in data quality issues?
11. Are reports widely announced and effectively disseminated to intended users?

### **Standard 8: Burden Reduction**

1. Is all the data that is requested actually used for either reporting to Congress, management improvements, or technical assistance within two years of collection?
2. Before requiring any data, was there a review of data already available being submitted by the same grantees through other federal programs?
3. Is there on-going communication with offices providing similar services or targeting similar customers/grantees?
4. Were grantees and other key stakeholders, such as states, included on the data collection decisions?