Best Practices for Determining Subgroup Size in Accountability Systems While Protecting Personally Identifiable Student Information

INSTITUTE OF EDUCATION SCIENCES CONGRESSIONALLY MANDATED REPORT

2018 Forum

Marilyn Seastrom

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

Congressional Charge to IES

Produce and widely disseminate a report on *"best practices for determining valid, reliable, and statistically significant minimum numbers of students for each of the subgroups of students"* (ESSA 2015)

A result is statistically valid if –

- it accurately measures what it is intended to measure;
- the result can be generalized to other places, people, and times; and
- the statistical conclusions drawn from the result are reasonable (i.e., credible or believable).

A result is statistically reliable if –

- it is consistent, stable, and reproducible from one use to the next,
- It is of high quality, and
- relatively error free.

Why is a minimum n-size needed?

- Minimum n-size refers to the lowest statistically defensible subgroup size that can be reported with protections for personally identifiable information in a state accountability system.
- The minimum n-size a state establishes and the privacy protections it implements will directly determine how much data will be publicly reported in the system.

To determine the minimum number of students for reporting

- Select a population perspective that treats the population measured as a universe and uses descriptive statistics to study the outcomes for a particular set of students, OR
- Select a sample perspective that treats the population measured as a sample from a larger population of similarly defined groups of students over time.

Population Perspective

1. How large a difference between two values must be to qualify as meaningful; that is how many percentage points of change are required for a meaningful difference?
2. How many students must have a change in status for the change to be recognized as a meaningful difference?
   - A difference of only x students should not produce a meaningful change.

Population Perspective

- What is the size of the reporting subgroups ?
  - A small change in a small subgroup can result in a relatively large change
- Goal: strike a balance between the number of students required to trigger a meaningful difference and the size of the smallest population that will yield such a meaningful difference.

Table 1. Percent of population represented by a change of 2, 3, 4, or 5 students at different population sizes.

	Small Number of Students					Small Number of Students			
Population Size	2	3	4	5	Population Size	2	3	4	5
5	40	60	80	100	60	3	5	7	8
6	33	50	67	83	61	3	5	7	8
7	29	43	57	71	62	3	5	6	8
8	25	38	50	63	63	3	5	6	8
9	22	33	44	56	64	3	5	6	8
10	20	30	40	50	65	3	5	6	8
11	18	27	36	45	66	3	5	6	8
12	17	25	33	42	67	3	4	6	7
13	15	23	31	38	68	3	4	6	7
14	14	21	29	36	69	3	4	6	7
15	13	20	27	33	70	3	4	6	7
16	13	19	25	31	71	3	4	6	7
17	12	18	24	29	72	3	4	6	7
18	11	17	22	28	73	3	4	5	7
19	11	16	21	26	74	3	4	5	7
20	10	15	20	25	75	3	4	5	7
21	10	14	19	24	76	3	4	5	7
22	9	14	18	23	77	3	4	5	6
23	9	13	17	22	78	3	4	5	6
24	8	13	17	21	79	3	4	5	6
25	8	12	16	20	80	3	4	5	6
26	8	12	15	19	81	2	4	5	6
27	7	11	15	19	82	2	4	5	6
28	7	11	14	18	83	2	4	5	6
29	7	10	14	17	84	2	4	5	6
30	7	10	13	17	85	2	4	5	6
31	6	10	13	16	86	2	3	5	6
32	6	9	13	16	87	2	3	5	6
33	6	9	12	15	88	2	3	5	6
34	6	9	12	15	89	2	3	4	6
35	6	9	11	14	90	2	3	4	6
36	6	8	11	14	91	2	3	4	5
37	5	8	11	14	92	2	3	4	5
38	5	8	11	13	93	2	3	4	5

Population Perspective Scenario

- Treat differences as meaningful only if they are greater than 10 percentage points.
- Do not allow a change of 3 students to trigger a meaningful difference.
- 3 out of 29 students would produce a 10 percentage point change, but would not produce a change greater than 10 percentage points—minimum n = 29

Sample Perspective

- Recall that a sample perspective treats the population measured as a sample from a larger population of similarly defined groups of students over time.
- Assuming the population parameters are unknown and must be estimated using methods of statistical inference, can statistically significant differences be detected for subgroups at the school, district and state levels?

Sample Perspective

At a specified level of confidence subgroups with smaller numbers of students will have larger margins of error, leading to questions:

- *Does the observed margin of error meet the state's criteria for statistical conclusion validity?*
- *Does it provide useful information about student progress?*
- *Do estimates meet the reproducibility, consistency, and stability criteria of reliability?*

Sample Perspective Scenario

- Assume: Level of confidence = 95% percent and 50% of students are at or above proficient

Level	# students in subgroup	Margin of error	Confidence interval size	Confidence interval range
School	30	17.9	35.8	32.1/67.9
District	190	7.1	14.2	42.9/57.1
State	3,500	1.7	3.4	48.3/51.7

Population vs Sample Perspective

Assume a meaningful or significant difference of more than 10 percentage points:

- for a subgroup change or difference of more than 10 percentage points that is not triggered by 3 or fewer students, **the minimum n-size is 29.**
- using a sampling approach to test for a difference of more than10 percentage points with 50 percent of the students in a subgroup in one category, **the minimum n-size is 96.**

Population vs Sample Perspective

Assume a meaningful or significant difference of more than 10%age points and 50% of students at or above proficient

		# and % at or above 50% for sig/diff			
		Population		Sample	
Level	# students	#	%	#	%
School	30	19	63.3	23	70.0
District	190	115	60.5	109	57.0
State	3,500	2,101	60.0	1,811	52.0

Population vs Sample Perspective

- Clear trade-off between the two perspectives
- IF margin of error **>** pre-established meaningful difference, the subgroup will demonstrate progress using a population perspective but not with a sampling perspective (i.e., small subgroups).
- But IF margin of error **<** pre-established meaningful difference, more larger subgroups will demonstrate progress using a sampling perspective

Population vs Sample Perspective

- With sampling perspective, smaller margin of errors at the state level = more significant at state level than district level
- Same pattern for district compared to schools.
- Level the playing field by defining differences as those differences that are statistically significant and meet the pre-established percentage point difference.

Protecting Data in Reporting

- Identify recommended privacy controls to be used (such as primary and complementary suppression, ranges, top and bottom coding, and rounding) to ensure that personally identifiable information is not inadvertently disclosed
- Confirm that the specified minimum number, in combination with the privacy controls, is sufficient to not reveal any personally identifiable information.

Protecting Data in Reporting

- Once a minimum number of students is selected, additional data protections are likely to be needed, BUT
- Each additional action taken to protect data for public release has a potentially negative impact on the remaining amount and quality of information available for reporting.
- Considering using data for policy decisions before adding protections (requires transparency).

Protecting Data in Reporting

- Even with a minimum n, displaying results for a small category within a subgroup can inadvertently lead to the identification of an individual student.
- Specify a minimum cell size required for reporting—threshold rule (usually 3 or 5).
- Suppress cells that fall below the threshold—primary
- Suppress only one category and it can easily be reconstructed—use complementary suppression.

Protecting Data in Reporting

- CAUTION: “While it is possible to select cells for complementary suppression manually, in all but the simplest of cases, it is difficult to guarantee that the result provides adequate protection” Federal Committee on Statistical Methodology Working Paper (page 17).
- One solution to this problem is to combine cell suppression with other data protection techniques.

Protecting Data in Reporting

- Use top and bottom coding to avoid reporting that all or nearly all (or none or nearly none) of the students in a population or subgroup share the same achievement level or the same outcome.
- Other parts of the distribution can be recoded into ranges or categories to reduce data loss that occurs with small cell suppression.
- Maximizes the amount of information that can be released.

Protecting Data in Reporting

Population Size	Reporting Ranges
0 - 5	Suppressed
6 - 15	<50%, ≥50%
16 - 30	≤20%, 21-39%, 40-59%, 60-79%, ≥80%
31 - 60	≤10%, 11-19%, 20-29%, 30-39%, 40-49%, 50-59%, 60-69%, 70-79%, 80-89%, ≥90%
61 - 300	≤5%, 5-9%, 10-14%, 15-19%, 20-24%, 25-29%, 30-34%, 35-39%, 40-44%, 45-49%, 50-54%, 55-59%, 60-64%, 65-69%, 70-74%, 75-79%, 80-84%, 85-89%, 90-95%, ≥95%
301 - 3,000	≤1%, whole number percentages, ≥99%
More than 3,000	≤0.1%, percentages to one decimal place, ≥99.9%

Protecting Data in Reporting

- The ED DRB has used a schema that includes reporting the group and subgroup totals and using a threshold of 3 for the reported ranges (i.e., each percentage in a displayed range could represent at least 3 students).

Protecting Data in Reporting

- Rounding refers to altering a number to another approximately similar value for the purpose of convenience or, in this context, to introduce an acceptable level of uncertainty that protects data values without substantially changing their meaning.
- For example, in a subset of the CRDC,
- values of 1, 2, and 3 are reported as 2;
- values of 4, 5, and 6 are reported as 5;
- etc.

Protecting Data in Reporting

- Given the hierarchical nature of education data, policymakers and analysts should understand that protection decisions made for one level of data (such as data intended to be used in schools) may limit the amount of detail that can be reported at another level (such as the district, state, or nationally).

Thank You!

Contact information:

Marilyn.Seastrom@ed.gov

Report Reference:

Seastrom, Marilyn (2017). Best Practices for Determining Subgroup Size in Accountability Systems While Protecting Personally Identifiable Student Information. (IES 2017-147). U.S. Department of Education, Institute of Education Sciences. Washington, DC. Retrieved [date] from http://ies.ed.gov/pubsearch.