

School Safety Dashboard: Data Privacy Considerations

The balance between transparency and data privacy

Public education dashboards and reports offer the challenge of meeting transparency requirements while also meeting legal requirements to protect each student's personally identifiable information (PII).

Even aggregate data can sometimes reveal information about individual students. This is particularly true when a member of a school community has knowledge that they can combine with public data to draw conclusions about individual students. Additionally, even individuals from outside the school community could combine information from multiple sources of public information to infer information about individual students—which is a risk increasingly made easier by advancements in AI and machine learning.

Introduction to disclosure avoidance

To prevent accidental disclosure of student information, the staff who create public-facing dashboards and reports apply what we call disclosure avoidance techniques to any dashboards and reports that are intended for the public. Disclosure avoidance techniques include strategies such as reporting at the state or LEA level and not at the school level, not reporting low counts of students, and reporting only ranges rather than actual percentages. Using these techniques helps protect students so that someone who has other information about a student—like a member of the school community—can't use the report to make inferences about individual students in the school.

For more information about disclosure avoidance and why it matters, please view the [Introduction to Disclosure Avoidance video](#).

Disclosure avoidance and the School Safety Dashboard

Reported values at the school level will likely be very low, which necessitates applying disclosure avoidance techniques to protect student anonymity. Therefore, many values at the school level will likely be suppressed, meaning that they will display as "n < 10" instead of the actual count.

USBE Disclosure Avoidance Policy

USBE's policy on disclosure avoidance is found below. This language is adapted from the policy on datagateway.schools.utah.gov.

Public education reports offer the challenge of meeting transparency requirements while also meeting legal requirements to protect each student's personally identifiable information (PII). Recognizing this, the

reporting requirements state that subgroup disaggregation of the data may not be published if the results would yield personally identifiable information about an individual student. While data used by the Utah State Board of Education (USBE) and local education agencies (LEAs) is comprehensive, the data made available to the public is masked to avoid unintended disclosure of personally identifiable information at summary school, LEA, or state-level reports.

This is done by applying the following statistical method for protecting PII.

1. Underlying counts for groups or subgroups totals are not reported.
2. If a reporting group has 1 or more subgroup(s) with 10 or fewer students.
 - The results of the subgroup(s) with 10 or fewer students are recoded as “N<10”
 - For remaining subgroups within the reporting group
 1. For subgroups with 300 or more students, apply the following suppression rules.
 1. Values of 99% to 100% are recoded to $\geq 99\%$
 2. Values of 0% to 1% are recoded to $\leq 1\%$
 2. For subgroups with 100 or more than but less than 300 students, apply the following suppression rules.
 1. Values of 98% to 100% are recoded to $\geq 98\%$
 2. Values of 0% to 2% are recoded to $\leq 2\%$
 3. For subgroups with 40 or more but less than 100 students, apply the following suppression rules.
 1. Values of 95% to 100% are recoded to $\geq 95\%$
 2. Values of 0% to 5% are recoded to $\leq 5\%$
 4. For subgroups with 20 or more but less than 40 students, apply the following suppression rules.
 1. Values of 90% to 100% are recoded to $\geq 90\%$
 2. Values of 0% to 10% are recoded to $\leq 10\%$
 3. Recode the percentage in all remaining categories in all groups into intervals as follows (11-19,20-29,...,80-89)
 5. For subgroups with 10 or more but less than 20 students, apply the following suppression rules.
 1. Values of 80% to 100% are recoded to $\geq 80\%$
 2. Values of 0% to 20% are recoded to $\leq 20\%$
 3. Recode the percentage in all remaining categories in all groups into intervals as follows (20-29,30-39,...,70-79)

Additional resources

- [Data De-identification: An Overview of Basic Terms](#)
- [Frequently Asked Questions - Disclosure Avoidance](#)