

Cradle-to-Career Data System

Suppression Protocol for Summary Data

The Cradle-to-Career System data suppression protocol is based on an extensive review of legal requirements and statistical methods that was conducted by the California Health & Human Services Agency regarding the federal Health Insurance Portability and Accountability Act (HIPAA) and California's Information Practices Act (IPA).¹ It also references position papers by the National Center for Education Statistics.² HIPAA and IPA requirements align with (and exceed) those required for Family Educational Rights and Privacy Act (FERPA).

For Cradle-to-Career System public-facing tools, including the dashboards and query builder, and for summary data released through the data request process, the following mechanisms will be used to protect the identity of individuals.

To facilitate the ability of non-researchers to make sense of the information, all figures will be expressed as whole numbers, and in all cases where figures are unsuppressed, counts will be included to contextualize rates.

In cases where access to anonymized, individual-level data are approved through the data request process, a Disclosure Review Board will review any data before they can be removed from the secure data enclave. The Disclosure Review Board will adhere to the policies of the data providers that are participating in the data request.

The minimum values for suppression are a preliminary recommendation that should be tested once the data system is built to ensure that the protocol provides sufficient protections for individual identities while providing usable data.

¹ California Health and Human Services (2016). *Data De-identification Guidelines*.

² National Center for Education Statistics. (2010, December). *Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting*. <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2011603> and National Center for Education Statistics. (2011, October). *Statistical Methods for Protecting Personally Identifiable Information in the Disclosure of Graduation Rates of First-Time, Full-Time Degree- or Certificate-Seeking Undergraduate Students by 2-Year Degree-Granting Institutions of Higher Education*. <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2012151>

- 1) Information will not be released for data points that include information on less than 30 people in the overall sample (overall denominator). In cases where information has been suppressed, the count of individuals will not be displayed.

Example where there are less than 30 people in the denominator for an overall figure

	<i>Percentage</i>	<i>Count</i>
<i>Among Middletown Unified District 2018 graduates who enrolled in UC San Diego and graduated with a bachelor's degree, the number who earned a degree in Biology</i>	<i>22%</i>	<i>2 / 9</i>

Query builder display:

	<i>Percentage</i>	<i>Count</i>	<i>Missing Information</i>
<i>Among Middletown Unified District 2018 graduates who enrolled in UC San Diego and graduated with a bachelor's degree, the number who earned a degree in Biology</i>			<i>Fewer than 30 people</i>

- 2) Information will not be released for data points that include information on less than 11 people in a specific category (category denominator), excluding cases where there are no people in that category. If there are more than zero but less than 11 people in any category, then the next largest value will be suppressed until there are at least 11 people in the pool of suppressed individuals. In cases where information has been suppressed, the count of individuals will not be displayed. The Not Reported category will be suppressed in cases where one or more other categories has been suppressed, unless all other categories are unsuppressed and there is no data that falls into the Not Reported category.

Example where there are no people in only the Not Reported category

Underlying data:

Graduates of Biology Programs Who Attained a Living Wage		
	<i>Percentage</i>	<i>Count</i>
<i>Foster Youth</i>	85%	11 / 13
<i>Not Foster Youth</i>	82%	241 / 294
<i>Not Reported</i>	--	0 / 0

Query builder display:

Graduates of Biology Programs Who Attained a Living Wage			
	<i>Percentage</i>	<i>Count</i>	<i>Missing Information</i>
<i>Foster Youth</i>	85%	11 / 13	
<i>Not Foster Youth</i>	82%	241 / 294	
<i>Not Reported</i>			Not applicable
<i>Total</i>	82%	252 / 307	

Example where there are less than 11 people in only one category

Underlying data:

Graduates of Biology Programs Who Attained a Living Wage		
	<i>Percentage</i>	<i>Count</i>
<i>Students with a Disability</i>	29%	2 / 7
<i>Students that Do Not Have a Disability</i>	84%	241 / 288
<i>Not Reported</i>	85%	11 / 13

Query builder display:

Graduates of Biology Programs Who Attained a Living Wage			
	<i>Percentage</i>	<i>Count</i>	<i>Missing Information</i>
<i>Students with a Disability</i>			Hidden to protect individual identities
<i>Students that Do Not Have a Disability</i>	84%	241 / 288	
<i>Not Reported</i>			Hidden to protect individual identities
<i>Total</i>	82%	254 / 308	

Example where there are less than 11 people in more than one category

Underlying data:

Graduates of Biology Programs Who Attained a Living Wage		
	<i>Percentage</i>	<i>Count</i>
<i>American Indian or Alaska Native</i>	--	0 / 0
<i>Asian</i>	78%	31 / 40
<i>Black or African American</i>	30%	3 / 10
<i>Hispanic or Latino</i>	56%	76 / 136
<i>Native Hawaiian or Other Pacific Islander</i>	100%	1 / 1
<i>White</i>	64%	77 / 120
<i>Two or More Races</i>	48%	11 / 23
<i>Not Reported</i>	60%	31 / 52

Query builder display:

Graduates of Biology Programs Who Attained a Living Wage			
	<i>Percentage</i>	<i>Count</i>	<i>Missing Information</i>
<i>American Indian or Alaska Native</i>	--	0	
<i>Asian</i>	78%	31 / 40	
<i>Black or African American</i>			Hidden to protect individual identities
<i>Hispanic or Latino</i>	56%	76 / 136	
<i>Native Hawaiian or Other Pacific Islander</i>			Hidden to protect individual identities
<i>White</i>	64%	77 / 120	
<i>Two or More Races</i>	48%	11 / 23	
<i>Not Reported</i>			Hidden to protect individual identities
<i>Total</i>	60%	230 / 382	

- 3) To facilitate access to data while protecting individual privacy, when results are suppressed, users of the query builder tool or the data request process will be prompted to view results at a less disaggregated level

Appropriate groupings should be developed as part of the user centered design process. Examples of possible options include:

- Multiple years of data, such as a three-year timeframe rather than a single year
- Group multiple sub-categories together, such as viewing race/ethnicity groupings of *White* and *Non-White* or age brackets of *Age 24 or Older* and *Under Age 24*