### Public-use Microdata

[*The public-use data may require certain perturbations in order to ensure respondent privacy:*

* *Top and bottom-coding of variable when outliers can be a source for disclosure risk;*
* *Variable suppression and variable collapsing – particularly of PII data. Certain variables, such as RACE, for example, may not include all of the breakouts of race, but only include the collapsed version;*
* *Entire files may be suppressed – particularly when the geographic area is known.*

*If an external public file exists that contains variables found in the survey data, it is necessary to analyze these data via probabilistic matching to determine whether a potential disclosure risk exists. This is normally conducted using approved probabilistic matching such as FRIL.*

*Public-use data release can include:*

* *Microdata;*
* *Data access via an approved data tool ;*
* *Tables and reports generated from the study data.*

*Unless the requirement is waived by the DRB, all survey data files from the study must undergo random data swapping. DataSwap is the approved software for conducting this procedure.*

***If Both Restricted-use and Public-use Microdata Dissemination:***

*Please refer to the treatment of restricted-use microdata. The public-use data are based on the restricted-use data that have undergone the SDC measures. The comparable variables between the two files must match for two reasons: (1) consistency in analysis; and, (2) differences could provide clues to users regarding SDC procedures.*

***If Public-use Only Dissemination***

*When a data release is public-use only, it assumes that there is no intention of providing any audience additional identifiers reserved for restricted-use.*

***STANDARD 4-2-8:*** *For* [*public-use data files*](https://nces.ed.gov/statprog/2002/glossary.asp#public-file)*, NCES minimizes the possibility of a user matching outliers or unique cases on the file with external (or auxiliary) data sources. Because public-use files allow direct access to individual records, perturbation and* [*coarsening disclosure limitation techniques*](https://nces.ed.gov/statprog/2002/glossary.asp#coarsening) *may both be required. The* [*perturbation disclosure limitation techniques*](https://nces.ed.gov/statprog/2002/glossary.asp#perturbation) *by definition, include the techniques applied in a* [*confidentiality edit*](https://nces.ed.gov/statprog/2002/glossary.asp#c-edits) *(if one is performed) and may include additional perturbation disclosure limitation techniques as well.*

***Table. Methods for Protecting Individually Identifiable Data***

|  |  |  |
| --- | --- | --- |
| ***Type of Protection*** | ***Methods*** | |
| ***Perturbation*** | ***Coarsening*** |
| *Confidentiality Edit* | *Yes* | *Yes* |
| *Disclosure Limitation Techniques* | *Yes* | *Yes* |

*All public-use files (i.e., the edited restricted-use files) that contain any potentially individually identifiable information must undergo a* [*disclosure risk analysis*](https://nces.ed.gov/statprog/2002/glossary.asp#disclosure) *in preparation for release to the public. The steps are as follows:*

1. *At an early stage in designing and conducting this analysis, staff must consult the Disclosure Review Board (DRB) for guidance on disclosure risk analysis and on the use of NCES disclosure risk software. Any modifications that are necessary as a result of the analysis must be made, and the entire process must be documented.*
2. *The documentation of the disclosure risk analysis must be submitted to the DRB. The documentation must include descriptions of the risk of disclosure and the types of edits used to avoid disclosure. Decisions over the type of* [*confidentiality edits*](https://nces.ed.gov/statprog/2002/glossary.asp#c-edits) *must take into account the procedures needed to avoid disclosure of individually identifiable information, age of the data, accessibility of external files, detail and specificity of the data, and reliability and completeness of any external files. The documentation should also include the results demonstrating the disclosure risk after adjustments to the data.*
3. *The DRB will review the disclosure risk analysis report and make a recommendation to the Commissioner of NCES about the file release.*

*The Commissioner then rules on the release of the data file*.]