Additional Details and Criterion for Special Data Requests (SDRs)

What are SDRs and when are they appropriate?

Special Data Requests (SDRs) are a novel user support service provided by the National Data Archive on Child Abuse and Neglect (NDACAN). SDRs are ad hoc tabulations of select datasets archived at NDACAN (i.e., NCANDS, AFCARS, and NYTD), which are compiled by NDACAN staff and delivered to requestors.

The SDR process is designed to deliver straightforward, well-defined general summaries and counts (at the state or national level) to individuals who don't have the statistical or programming resources to compile the results from the full dataset(s) on their own.

The applications are reviewed by NDACAN staff and experts at the Children's Bureau, Administration for Children, Youth and Families, U.S. Department of Health and Human Services. Once an SDR application is approved, NDACAN staff compile the table(s) utilizing the relevant data and the tables are reviewed and approved by the Children's Bureau prior to submission to the requestors.

There are many instances where SDRs are not appropriate or sufficient for individuals. An SDR would *not* be appropriate for:
- Those who plan to use the data for multiple and/or long-term projects;
- In-depth or advanced analyses or modeling;
- Thesis or academic projects where the NDACAN data requested is the primary source of information;
- Those who want to cross-examine the interaction of a multitude of variables (i.e., slicing the data many ways);
- Those who have proportionately small or distinctive populations of interest.

How do you apply for full dataset(s)?

For cases where an SDR is not appropriate, individuals should apply for access to the full dataset(s) of interest from NDACAN, as detailed here:
- For general release datasets (e.g. AFCARS, NYTD):
  https://www.ndacan.acf.hhs.gov/datasets/request-dataset.cfm
- For or restricted release datasets, which require an IRB (e.g. NCANDS):
  https://www.ndacan.acf.hhs.gov/datasets/request-restricted-data.cfm

Contact NDACANsupport@cornell.edu for more information about ordering dataset(s).

What are other resources to check before submitting an SDR?

Before submitting a SDR application to NDACAN, individuals should review the existing publicly available information and tables for the large administrative datasets as follows:

- For NCANDS:
  - The Child Maltreatment Reports (https://acf.gov/cb/data-research/child-maltreatment)
- For AFCARS:
  - The AFCARS Reports (https://acf.gov/cb/research-data-technology/statistics-research/afcars)
  - The AFCARS Dashboard (https://tableau-public.acf.gov/views/afcars_dashboard_main_page/mainpage?%3Aembed=y&%3AisGuestRedirectFromVizportal=y)
- For NCANDS or AFCARS:
  - Child Trends data (https://www.childtrends.org/publications/state-level-data-for-understanding-child-welfare-in-the-united-states)
  - The Annie E. Casey KIDS COUNT Data Center (https://datacenter.aecf.org/)
- For NYTD:
  - NYTD Data Briefs and Data Snapshots (https://acf.gov/cb/data-research/data-and-statistics-nytd)

Contact NDACANsupport@cornell.edu if you are unsure whether the SDR process is appropriate for your needs.

What are common reasons that SDRs are rejected?

An application is not guaranteed to be approved once submitted. There are many factors that go into the review and approval of an application, and requests can be denied for many reasons, such as those listed below.

1. If the request asks for information that is not in the public release files, or requests that violate the suppression and masking rules applied to the public release files (e.g., no counties with less than 700 records, no unmasked dates or variables).
   - **Why**: This information is not available to any requestor/data user through NDACAN.
   - **What to do**: You should revise your request and/or research strategy to avoid using any masked or suppressed data elements.

2. If the request threatens data confidentiality or has high disclosure risks, such as requests leading to a large proportion of small cell counts (i.e., less than 10).
   - **Why**: It is NDACAN policy to suppress small counts between 1-9. If a request will result in a large proportion of suppressed cells, then the interpretability, reliability, and utility of the tables will be too severely affected. Depending on the variables(s), NDACAN may also choose to suppress cells with 0s.
   - **What to do**: Usually in such requests, the data are being cross tabbed on too many variables or filters (i.e., the scope is too narrow, or one or more variables are poorly reported/populated). You should broaden the scope of your request, possibly by regrouping variable levels, dropping variables or filters, choosing comparable variables that are more reliable, and/or utilizing multiple variables in combination.

Alternatively, these requests may be better suited for ordering the full dataset(s) and investigating the context that leads to such small counts.

3. If the request is overly complex or time-consuming.
    o **Why**: Requests are deemed "overly complex" or "time-consuming" if they, for example, rely on poorly reported variables, include too many filters, or yields an overly abundant amount of output (e.g. very many columns, numerous large tables). "Complex" requests almost always result in the outcome outlined above in (2.). Constructing a complexly defined table takes many thoughtful steps and decisions that may affect final counts, such as choosing how and in what order to filter and aggregate variables across different dimensions (e.g. time, geography, demographics), creating intermediate variables, or resolving missing or inconsistent values. Every additional layer of complexity included in a request requires additional work and review from staff at NDACAN and the Children's Bureau. Lastly, requests such as these suffer from issues of losing broader data context, and poor interpretability and comparability.
    o **What to do**: You should revise your request following similar guidance outlined above. However, if your research question relies on such level of analyses, you should order the dataset(s) directly and investigate more fully the context of the data and outcomes.

4. If the request includes variables of highly irregular or poor reporting patterns.
    o **Why**: The quality and consistency of variables varies greatly: by year and geographic area for which laws and policies are bounded; by demographic characteristics such as race, sex, or age; and by other factors directly or not directly measured. Interpretations and comparisons with such tables may not be valid or meaningful without the full data context or additional analyses to understand such reporting mechanisms. Additionally, requests such as these usually lead to problems outlined above.
    o **What to do**: You should revise your request and/or research plan to avoid using such variables. You can narrow the scope of your tables to years and/or states where reporting is reliable, if possible. You can try to identify other related variables that are more reliable to use instead, similar to the guidance outlined above. You can also order the dataset(s) directly and investigate more fully the context of the data and outcomes.

5. If the requestor has submitted more requests than permitted.
    o **Why**: Applicants may request no more than six data tables per form. Requestors may only have one SDR filled every few months, unless it is a continuation of a prior request. This relates to the exclusion criteria of overly time-consuming requests. NDACAN staff are not meant to fulfill the function of a private statistical consultant or researcher, and staff cannot provide inordinate time to individual requestors.
    o **What to do**: You should order the dataset(s) to directly utilize in your research. Requests for a multitude of tables indicates that the research goal requires more detailed and advanced analyses than SDRs are meant to provide.