**NATS Small Cell Suppression Implementation**

**June 2022**

**Rules**

1. **Primary suppression (N)**: Suppress **any/all N** (number of participants, for instance) that is between 1 and 5 (inclusive).  If a number (N) is suppressed, then the corresponding percentage (and other statistics related to N) should also be suppressed.  [If the report does not display N but only related statistics then suppression is still based strictly on the hidden, underlying N.]
2. **Secondary suppression (N)**: **If** there is a **Total** row (and/or column) to which all the **detail data sums** then:

**If** only **one N** from #1 was suppressed

**or**

(more than one N from #1 was suppressed and the **sum of all N** suppressed in #1 between 1 and 5 (inclusive))

then:

also suppress the next lowest N (if another N exists>5)

1. [possible improvement – rare condition] If a table requires both row and column secondary suppression and the total value has already been suppressed by one or the other, then no further secondary suppression is required. See tables 7 & 10 for possible examples. This would result in fewer cells being suppressed.

**General Macros**

The small cell suppression was implemented with three general, reusable macros. They assume

* The data are already aggregated in tabular SAS dataset (i.e. the macros evaluate the aggregated tables for small cell suppression).
* Primary suppression is run first for each table followed by secondary suppression, as needed, in either order.
* **%SuppressPrimary:** Performs primary suppression on the entire table, evaluating and suppressing based on the groups of variables in the VarList argument.

Parameters:

* Data - Input dataset, returns &Data
* Varlist - of the form N1 Stat1 Stat2|N2 Stat1 Stat2 (etc.)
* SuppressUpperBound: Upper bound needing suppression (default=5)
* SuppressValue - value to assign to supressed cells (default=-1)
* **%SuppressSecondaryRow:** Cell suppression of a single row of data. [Do not include the Total cell. It does not matter what the Total is, but only invoked if there is a total to which the sum of detail records in the row total.]

Parameters:

* Data: Input dataset, returns &Data.
* VarList: A pipe delimited list of groups of variables of the form N1 Stat1 Stat2|N2 Stat1 Stat2 (etc.).
* ObsList: The observation number from &Data that should be evaluated for small cell size.
* SuppressUpperBound: Upper bound needing suppression (default=5)
* SuppressValue - value to assign to supressed cells (default=-1()
* **%SuppressSeconaryCol:** Cell suppression of a column of data. [Do not include the Total cell. It does not matter what the Total is, but only invoked if there is a total to which the sum of detail records in the column total.]

Parameters:

* Data: Input dataset, returns &Data.
* VarList: A single group of variables of the form N1 Stat1 Stat2, where first (N) is to be evaluated for suppression in which case the rest of the are also suppressed.
* ObsList: Pipe-delimited list of observation numbers from &Data that should be evaluated for small cell size.
* SuppressUpperBound: Upper bound needing suppression (default=5)
* SuppressValue - value to assign to suppressed cells (default=-1)

**NATS Implementation Approach**

Both sets of tables are produced in a single run

* The full (unsuppressed) tables are produced as originally programmed.
* Cell -suppressed tables are produced with an “\_S” suffix to the name

Cell suppression will only be applied to the individual state reports, not the ALL-States report.

General macros: Assume that the data has already been aggregated into a SAS dataset. The macros suppress values in the aggregated dataset as per the general rules.

Suppressed values set to [configurable] -1.

If both Row and Column suppression needed (for NATS, tables 7 & 10): called Row suppression first, then column suppression; then need to invoke row suppression again to correct any column suppression that left a single row value suppressed.

Possible adjustment: exclude Missing/Not-Reported variables from primary and/or secondary suppression. No change to the generalized macros needed. The macro calls can be customized to exclude some variables from evaluation.

**Approach for each Table**

Table 1

Primary and secondary suppression.

‘Participant Category’ section (secondary suppression **not** needed): has Totals but as a practical matter, none of the N would be less than 5

‘**Age at Certification’** section: Women, infants, and children each has a Total that would require secondary **column** suppression

‘**Trimester of Certification**’ needs secondary **column** suppression: Total for ‘Pregnant Women’

‘**Race’** section requires secondary **column** suppression

‘**Ethnicity’** section requires secondary **column** suppression

‘**Income** as a Percentage’ secondary **column** suppression

‘**Number of People in Economic Unit’** secondary **column** suppression

‘**Number of Assigned Nutritional Ris**ks’ secondary **column** suppresion

Tables 2-6

Primary suppression only. No ‘Total’ Row – sum of the Nutrisks are not intended to add to the Total row.

Tables 7 & 10

Primary and Secondary suppression:

Each **row** has a total cell

each **column** has a Total Participants cell separate for each file

Tables 8 & 11

Primary and Secondary suppression:

each **column** has a Total Participants cell separate for each file

Tables 9 & 12

Primary suppression only:

Table 13a, 13b, and 13c

Primary suppression only

Tables 14-16

Note: In fact, because these reports are on the **top 5** nutritional risks, none of the rows would be small enough to require suppression. The code is in place, nonetheless.

Primary and secondary **row** suppression

Each **row** has a Total Participants cell.

Tables 18a – 24b

Primary suppression only