How Data De-duplication is handled within DATIM

Scope: This article describes the "De-Duplication Mechanism" that is used in the DATIM De- Duplication Application.

Related Guidance:

- DATIM De-Duplication Application User Guide
    - Describes how to use the DATIM De-Duplication Application
    - Does not incorporate guidance on how de-duplication fits within the PEPFAR business process.

The De-duplication Mechanism

DATIM has a single specialized Implementing Mechanism per Operating Unit, called the 'De- duplication Mechanism,' which is not assigned to an Implementing Partner. The De-duplication Mechanism is used to de-duplicate any given group of the same indicators occurring at the same site.

The De-duplication Mechanism stores a negative value for a given group of the same indicators occurring at the same site and the value is the amount by which the PEPFAR Country Team has determined should be decremented from the total that would be arrived at if the group of these values were simply summed.

How data are entered on the De-duplication Mechanism

There are two ways in which values are entered into the De-duplication Mechanism:

(a) Duplicates within a given Indicator/Site pair

This type of duplicate is presented to the user via the Data Deduplication Application and the data values are derived from the user selection made via that user interface.

(b) Duplicates between DSD and TA values within a given Indicator/Site pair

Where data values have been entered for both a DSD Indicator and for the same Indicator's TA component, the De-duplication Mechanism is automatically assigned the negative of the TA number for that Indicator/Site pair. Because you cannot have DSD and TA values for the same Mechanism/Indicator/Site, this situation only occurs when two mechanisms have entered data for the same Indicator/Site pair. In such cases, PEPFAR guidance requires that the DSD value be counted, and the TA value be ignored.

Viewing the De-duplication Mechanism

- Required User Rights

The De-duplication Mechanism is used by – and available to – DATIM users with 'Interagency' privileges. These rights can be assigned to any account created by the in-country DATIM User Administrator. In general, Interagency accounts are provided as 'secondary accounts' whenassigned to Agency users, as Interagency accounts are used for cross-agency functional support including things like data deduplication.

Accessing the Mechanism

The De-duplication Mechanism can be viewed by users with appropriate credentials via the Data Set Report under the Reports App. That report allows users to select outputs in one of three ways:

- Default output – this view calculates the values displayed by summing all values selected in the criteria box shown below, including the negative values stored in the De-duplication Mechanism

- 'All Mechanisms without deduplication' – this view is the same as the above, but does not include the values stored in the De-duplication Mechanism

- 'Deduplication adjustments' – this view presents the De-duplication Mechanism exclusively. Note that this is also the only way to review the

cancelled values for TA where they are excluded due to the presence of a DSD value for the same Indicator/Site pair, as discussed in 2.b above.

How the De-duplication Mechanism is used

The De-duplication Mechanism is used in DATIM as well as FACTS Info. Each of these is discussed below:

- In DATIM

In addition to the Data Set Report, discussed previously, the De-duplication Mechanism is incorporated in all analytic outputs (Data Visualizer and Pivot Table apps) when run with the 'Interagency' role and when the Implementing Mechanism tab is in default mode – meaning all Mechanisms are selected, rather than a specific subset.

- In FACTS Info

The De-Duplication Mechanism is also provided to FACTS Info as part of the nightly upload. The FACTS Info COP tab 'Indicators/Technical Area Level' table is populated based on the de-duplicated data values provided by DATIM. Technical Area Level Indicators are no longer manually entered into FACTS Info. Instead, they are generated based upon the values entered into DATIM, including the negative De-duplication Mechanism.