Generating Data Sets

Guidance for the Patient Safety Component

August 2024

Generating data sets is the first step to performing analysis in NHSN. This process will freeze your NHSN data at a specific point in time and copy those data into defined data sets. When you wish to see updates to the data contained in your NHSN reports, you must regenerate datasets for those updates to be reflected. Data sets are user-specific; therefore, each user in NHSN who wishes to analyze data must generate data sets.

This guide provides instructions for generating data sets in the NHSN Patient Safety (PS) Component, using a feature called "Timeboxing".

Instructions

1. Within NHSN, navigate to Analysis → Generate Data sets.
2. On the "Generate Data Sets" screen, the "Reporting Data Sets" tab will open.
   A. A time period can be specified, to apply to all analysis data sets.
   B. The "Generate Reporting Data Sets" button will start the data set generation process.
   C. The "Last Generated" information will tell you when your data sets were last generated, and for which time period. This information will also appear in the footnotes on all reports.

3. Specify the desired time period for your data sets, as illustrated in Step A above. The time period must be specified using month and year, however, there's additional guidance depending on the circumstance:
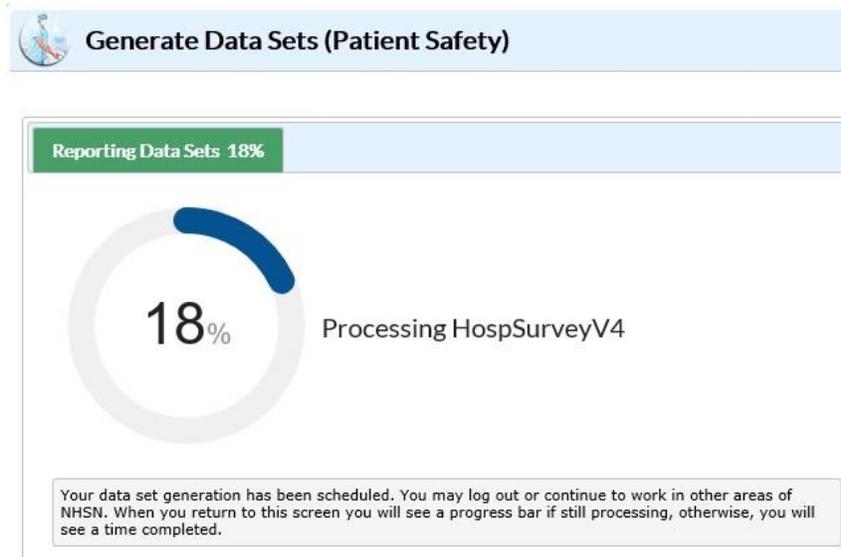
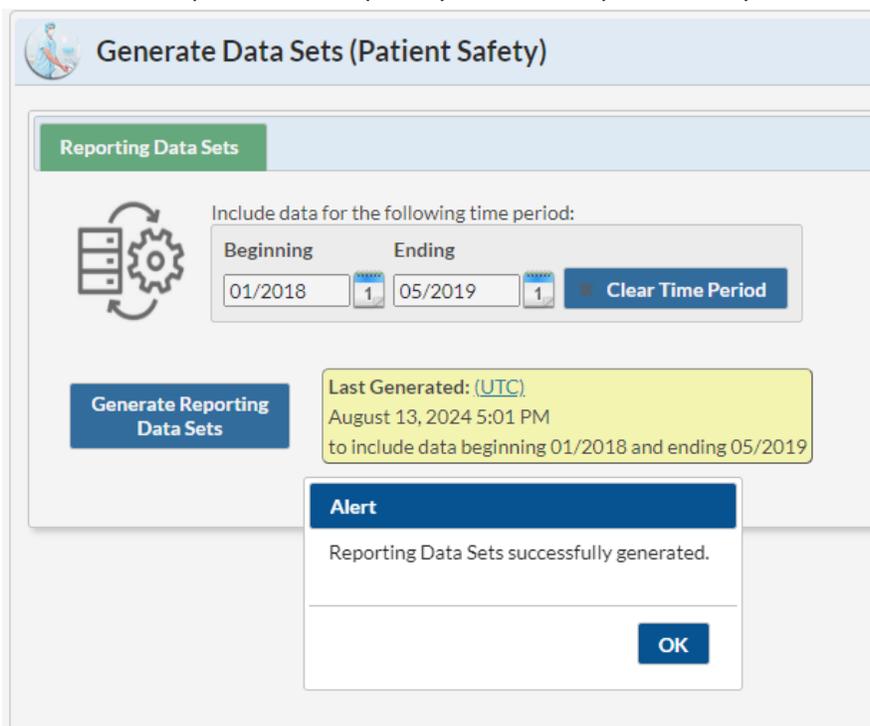|	
Include data for the following time period: **Beginning** `01/2019` **Ending** `03/2019` — **Clear Time Period**	**Both** beginning and ending = **non**blank Will include all data within the time period specified (for example, 01/01/19 thru 03/31/19)
Include data for the following time period: **Beginning** `01/2016` **Ending** `mm/yyyy` — **Clear Time Period**	Ending date = blank Will include all data from beginning date specified (01/2016) through present day
Include data for the following time period: **Beginning** `mm/yyyy` **Ending** `12/2018` — **Clear Time Period**	Beginning date = blank Will include all data through ending date specified (12/2018)
Include data for the following time period: **Beginning** `mm/yyyy` **Ending** `mm/yyyy` — **Clear Time Period**	**Both** beginning and ending = blank Will include all data reported, regardless of time period

Additional Notes:

- Survey data will **not** be limited by time period; all surveys completed by the facility will be included, regardless of survey year.
- SSIs with event dates up to 90 days past the "ending" date will be included.
  - *90 days is the maximum surveillance period for several operative procedure categories*
- The data set generation process will include all LabID events. Once the algorithms are applied for assignment on onset and CDI Assay (CDI only), the events outside the specified time period will be removed from the final analysis data sets.

4. Once the time period is specified, click "Generate Reporting Data Sets". A pop-up message will appear, asking you to confirm that you want the current data sets to be updated. Click "OK".

5. When the process starts, you'll see a new progress image appear on the Reporting Data Sets tab. At this point, you can navigate to other areas of NHSN, or log off. NOTE: You <u>cannot</u> run analysis reports while datasets are generating.

6. When data sets have generated, you'll receive an alert that the data sets generated successfully. In addition, the date and time Last Generated will be updated. At this point, you can run reports in Analysis.

Additional Resources:

Analysis Quick Reference Guides: https://www.cdc.gov/nhsn/ps-analysis-resources/reference-guides.html

Analysis Training: https://www.cdc.gov/nhsn/training/analysis/index.html

NHSN Helpdesk: nhsn@cdc.gov