# Community Health Status Indicators GIS Analyst
# User Documentation

Updated: October 8, 2009

**CHSI GIS Analyst Purpose**
The Community Health Status Indicators (CHSI) GIS Analyst is a web-based application that provides rich visualization of the community health status indicators that are included in the core CHSI data set. For more information on CHSI, please see http://communityhealth.hhs.gov.

**CHSI Indicator Types**
CHSI data are in a variety of forms:

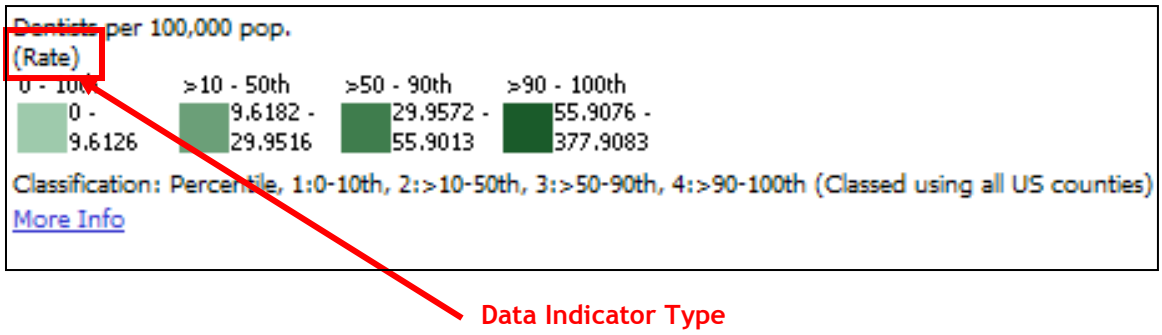
1. Rate: Rate-type CHSI measures typically represent the numbers of an element of interest that occur per a standard base measure such as population or area. For example, the "Access to Care/Dentists per 100,000" indicator measures the number of dentists per 100,000 people.

2. Boolean: Boolean-type (Yes/No) CHSI measures typically represent whether or not an item of interest is present. For example, the "Access to Care/Community/Migrant Health Centers" indicator captures the presence of community/migrant health centers in a county.

3. Count: Count-type CHSI measures typically represent the raw count of an element of interest that occurs in an area. For example, the "Access to Care/Disabled" indicator captures the count of disabled persons in a county.

4. Ratio: Ratio-type CHSI measures typically represent the ratio of an element of interest to another element of interest. For example, the "Environmental Health/Salmonella (Reported/Expected)" indicator captures the ratio of reported salmonella cases to expected salmonella cases (per county).

*Using the Interface*
The CHSI data indicator type is indicated above each map on the CHSI GIS Analyst interface.



**CHSI Percentile Classification**
All CHSI maps (except those using a Boolean-type data) are classed using the percentile-rank method. The percentile-rank of an observation is the percent of observations that are lower than that observation of interest. Thus, an observation with a percentile-rank of 65 (or "in the 65th percentile") is above 65% of the observations. In the case of the CHSI data, each observation was given a percentile-rank value using the formula:

$$\frac{cf_l + .5(f_i)}{N} \times 100\%$$

where $cf_l$ is the cumulative frequency for all scores lower than the score of interest, $f_i$ is the frequency of the score of interest, and N is the number of data points in the sample.

Thus, if a county's percentile-rank for an indicator was 48.6, then 48.6% of the observations for that indicator are below the indicator value for the county of interest.

The four classes were chosen in an effort to visually differentiate the counties with the lowest and highest values (for the indicator of interest) from the counties in the low-middle and high-middle ranges. The four classes, with an explanation for each, appear below:

The four classes chosen are:

1. 0 – 10th percentile: Represents the lower 10% of observations in a ranked list.

2. >10th – <50th percentile: Represents observations greater than the lower 10% and less than the top 50% (in a ranked list).

3. 50th - < 90th percentile: Represents observations in the upper 50% observations but lower than the upper 10% (in a ranked list).

4. 90th – 100th percentile: Represents the upper 10% of observations in a ranked list.

*Handling of Multiple Observations of the Same Value*
On occasion, multiple observations of the same data value do occur in the CHSI indicator data sets. In the event that a tie occurs, it is possible that one or more classes are not represented by any observations. Consider the following data set:

| Measurement Value | RANK | LOWER | TIES | Percent rank |
|---|---|---|---|---|
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 1 | 1 | 0 | 13 | 40.625% |
| 2 | 14 | 13 | 3 | 90.625% |
| 2 | 14 | 13 | 3 | 90.625% |
| 2 | 14 | 13 | 3 | 90.625% |
| | | | | |

The percentile rank of the “1” value is calculated thus:

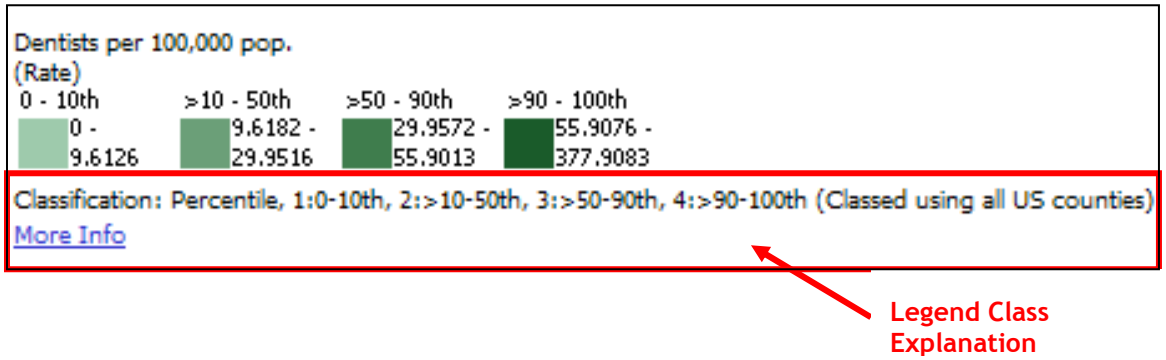
(0 + .5(13))/16 * 100%

6.5/16 * 100% =

0.40625 * 100 = 40.625

As you can see, because the value 1 appears in the dataset 13 times, the percentile-rank of 1 is equal to 40.625%. In the case of this data set, Class 1 and Class 3 would not be represented by any values.

*Using the Interface*
A brief explanation of each legend class appears below the legend.




*Universe of Observations*
The universe of observations used in the percentile ranking process includes observations for all counties in the United States. Thus, a county in the "90th – 100th percentile" range is in the top ten percent of *all* US counties for the indicator of interest.

**Handling of Missing Data or No Reported Data Values**
Due to a variety of reasons, CHSI data measures are occasionally missing, not reported, or otherwise unavailable. The classifications of these missing or not reported values appears below:

1. No Report, Too Few Events

   Explanation: No report is available for this county and indicator because a small number of events occurred during the specified time period.

   Example: See footnotes at:
   http://communityhealth.hhs.gov/MeasuresOfBirthAndDeath.aspx?GeogCD=13053&PeerStrat=39&state=Georgia&county=Chattahoochee

2. No Data Available for This County and Indicator

   Explanation: No data is available for this county and indicator. A database record does exist for this county and indicator and the record is coded as "No Data Available."

   Example: See footnotes at:
   http://communityhealth.hhs.gov/MeasuresOfBirthAndDeath.aspx?GeogCD=13053&PeerStrat=39&state=Georgia&county=Chattahoochee

3. Release of Data for All Counties Has Not Been Authorized (RNA)

   Explanation: No record is available for any county and indicator. Unlike the "No Data Available" (2) RNA represents a variable missing from the CHSI database. This is a result of the desired county-level data not being released.

4. Not Calculable (used only in the CHSI mapping facility.)

   Explanation: When the expected number of cases is zero the division of the observed by the expected results is an invalid operation.

**Comments**

Comments regarding the CHSI GIS Analyst and documentation are always welcome. Please respond to comments@hrsa.gov.