Tips for Examining Data

Comparisons

Comparing your data to other existing data provides a context for understanding your assessment results. For example, survey data from a high school may seem, at a glance, to reveal high smoking rates among 9th graders. However, a comparison of these data to statewide data might actually show that the school’s smoking rates are much lower than the statewide average.

When examining data, it is useful to make comparisons in some of the following ways:

- Between the community now and sometime in the past
- Between the community and the state (or similar communities)
- Among different population groups in the community, including different age groups and genders

For example, you can look at data from an indicator such as 30-day alcohol use and draw comparisons in different ways. For instance, you can compare it to the past and discover that your community has a lower or higher rate than it did previously. Or when comparing it to the state, you might discover that your community has a lower or higher rate than your state. When you compare population groups, you may notice whether there is a higher rate among males than females, or a lower rate among 14 year olds than among 17 year olds.

When looking at data and making comparisons, be careful using small numbers to calculate rates - they may cause rates to appear exaggerated.

Rates

In epidemiology, rates are a measure of the frequency with which a health event occurs in a specific population over a period of time. Rates are used to standardize data in order to be able to compare it across different population sizes.

For example, a town with a population of 30,000 has 500 arrests in a given year for driving under the influence (DUIs). Divide 500 by 30,000 to get a rate of 0.017 arrests per person per year. To make the rate easier to understand, multiply by 1,000 – now you can say that the rate of DUIs for the town is 17 per 1,000 people per year.

The following calculation provides a rate per 1,000 people per time period:

\[
\text{Rate} = \frac{\text{Number of cases}}{\text{Population over time period}} \times 1,000
\]

For national data sources or larger population sizes, the rate is often calculated with 10,000 or 100,000 population size. You can multiply by whatever size makes sense for your community.
Small numbers

When rates or percentages are calculated using small numbers, they may appear more exaggerated. For example, a 100% increase in motor vehicle accidents from 2003 to 2006 in Smithtown would sound like a lot of accidents. However, what if there had been one motor vehicle accident in 2003 and two in 2006? Although still a 100% increase, the severity is greatly diminished.

Small numbers can sometimes be misleading, as this example illustrates, when they are used to calculate percentages. A small number as a percentage may lead us to believe that there has been a significant increase, but when we look at the actual numbers we see that the increase is very small. Low, or even very low, percentages of some behaviors may be significant, however, and should not be dismissed as unimportant, especially if the behavior has severe consequences.

Things to Remember
- Examine different kinds of data – Substance use and other behavioral health problems are complex, so understanding them requires looking at different kinds of data, both quantitative and qualitative, to get an accurate and complete picture of the problems.
- Look for relationships and patterns – Numbers alone have no meaning, so look for patterns over time, as well as relationships between data.
- Notice any data gaps – After determining what data exists, you may discover gaps in the data or that you need additional data to answer certain assessment questions.
- Be aware that not all data are equal – Some data are more valid and available than others. You may need to use qualitative data to fill in some of your data gaps. Keep in mind that these data are less objective.

Comparison is one of the criteria when determining which problem to address. Other criteria can include:
- **Magnitude**: Which problem seems to be largest? (Beware small numbers.)
- **Time trend**: Is the problem getting worse over time?
- **Severity**: Is the problem resulting in mortality? Is it more costly?

**Make sure you collect data on these criteria.**