

## GRAPHING AVERAGES IN TUVA

April 2019, Molly Schauffler

Are your students wondering how to graph averages in Tuva? Your students might be asking: *What is the average amount of arsenic in wells in the New Hampshire schools?* To answer the question, they might be picturing a bar graph showing the average arsenic concentration for all of the wells from each school (Figure 1).

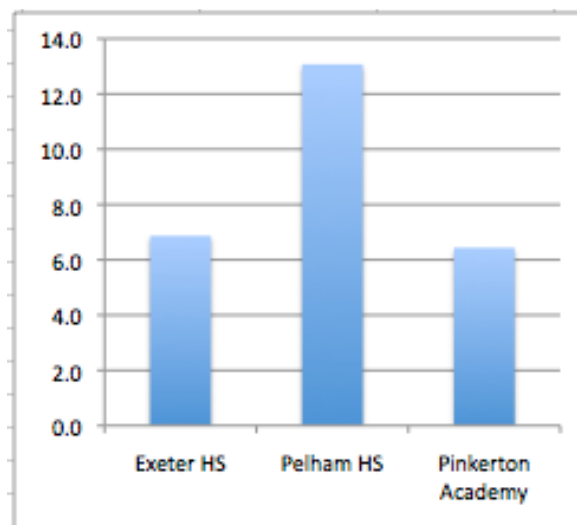


Figure 1 Arsenic (ug/L) in well samples (Excel)

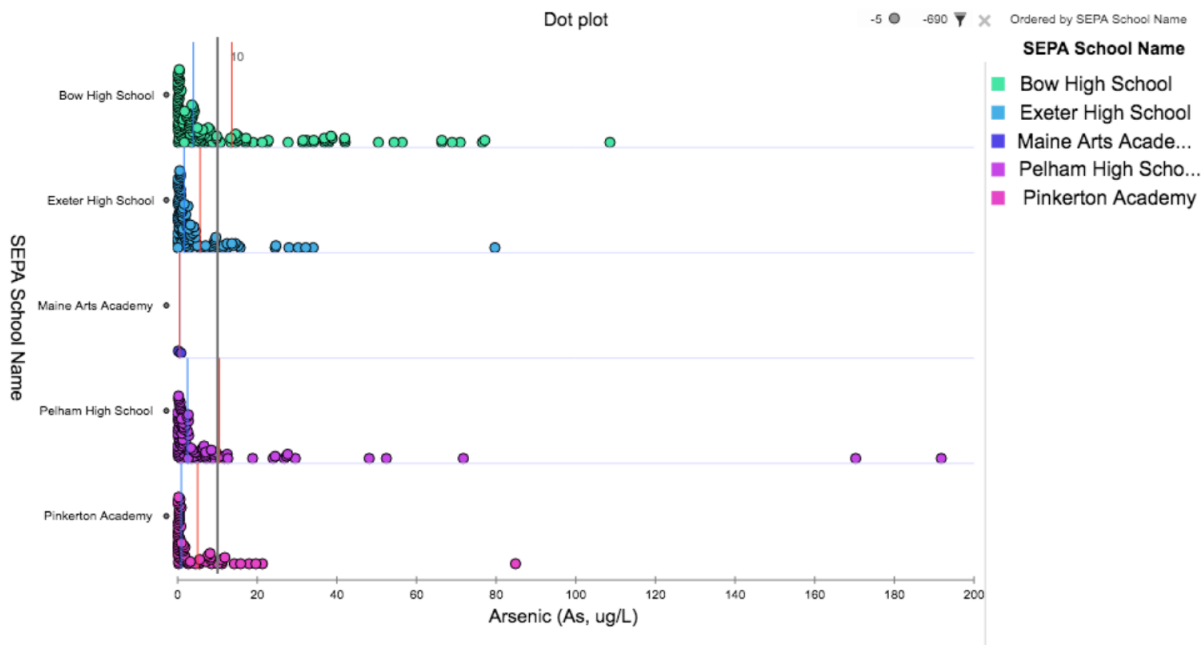
The graph answers the question, but this way of looking at the data doesn't tell very much about (1) how many wells were sampled at each school, or (2) what the distributions of arsenic concentrations were. In other words, did all the wells have concentrations close to average? Or were the concentrations wide ranging? Is average even a good way to represent all of the wells from each school?

The underlying question is really about *comparing the three schools in how much arsenic* is in the sampled wells. Were the wells similar or were they different? Finding an average is just one way to compare them, but other aspects of the samples tell much more of the story.

To compare the schools in Tuva, plot the distribution of data from each school as a dot plot. (To simplify, filter the data for just a few schools: click the pencil icon beside State, and uncheck all except New Hampshire). Put Arsenic on X and SEPA School Name on Y. Add the mean and the median by clicking on Stats, then Mean (red line), then Median (blue line). Students can add a reference line on Y to show the threshold for safe drinking water to help them understand what the comparisons mean in the context of the problem. (Click Stats, Add Reference Line to X, then position the line at 10 ug/L).

# GRAPHING AVERAGES IN TUVA

April 2019, Molly Schauffler



Now there is much more for students to talk about when comparing the schools. The distributions are strongly skewed, which means the median more fairly represents the center of the distribution than the mean does. The medians are quite different, and at two of the schools the medians are at or above the 10 ug/L. Most wells are below the threshold. The Maine Arts Academy had very few samples, and maybe not enough for a valid comparison.

Students have an enduring habit of summarizing data as group averages and plotting the averages in bar graphs. They will need coaching to look at data differently, as a distribution. When comparing two or more groups, encourage them to talk about more than just the average. They can compare the number of samples in each group, distribution shapes, ranges, the presence of extreme values, centers, how much the ranges overlap (and/or confidence intervals). They can talk about uncertainty. The bar graphs don't help you answer: *How sure can you be?*