The German Tank Problem—Activity Overview

During WWII Spies were asked to estimate the number of tanks the Germans had of various types. Allies were able to capture a number of German tanks and use a part of the serial number to estimate the number of tanks. The mathematicians came up with estimates quite a bit lower than those given by the spies. Long after the war it was discovered that spies had been deceived by the Germans repainting their tanks to increase their apparent numbers. The mathematicians were much closer to getting the right number of tanks.

To begin the activity, divide the students into two groups. Have one group use the following sample obtained by spies to estimate the number of tanks:

"Tank you very much!"
During World War II, in the European theatre, the commanders of the Allied forces noticed that German tanks had numbers stenciled onto them, and they began to suspect that the numbers were sequential; i.e. that the tank with the number 127 was the 127th tank to be built. Working on that premise, they attempted to determine how many tanks the German army had in the European theatre. Selecting a random sample of tanks yielded the following numbers:

236 489 392 12 167 263 469 172 245 254

How many tanks did the German army have?

The other group of students will take on the role of the mathematicians and try to estimate the number of German tanks using the following sample:

69, 314, 117, 16, 251,139, 177, 51, 311, 323

Tell your students the history of the problem and explain that this is how intelligence was gathered before satellites and technology. Each group is asked to come up with an estimate of N and describe the process they used to come up with their estimate. After about 10 minutes ask the students to explain their method (put the methods on the board with the group's estimate of the population size—group estimates as spies vs mathematicians)***Spies should be close to 500, Mathematicians should be close to 342.***

Common methods might include:	Examples of estimates of population size
Double the mean	358, 480, 404
Six times the standard deviation	874
Four times the standard deviation	515, 353
Sample max plus sample min	320

Third quartile plus one standard deviation	499
Double the median	408,644,212

Tell the class which one is the best estimator for each group and then ask them if that method is the best all the time.(The Max plus the Min is the best estimator.) Hopefully most students will see that the goodness of the estimator cannot be judged by a single estimate basked on one random sample. , that group may have gotten lucky.

Lead the students in a discussion about how one might judge estimators. The student s should arrive at the conclusion that many random samples should lead us to the best estimator. We will use simulations.

Students should then simulate their own random samples of size 10 (same as before) from a population that is known (use N = 500 or 342). Have the student groups create 50 simulated samples from the population and apply their method to each sample recording the estimate that each produces. When they're finished they should make a histogram of their estimates (this is the estimated sampling distribution of their statistic). Have the groups send the graphs to the teacher computer and place all the sampling distributions on the board. This is a good time to discuss sampling variability and bias. At this point we should be able to determine which method is best.

N = 342 is the right answer. In real life you wouldn't know what the right answer was. However you could perform simulation for a variety of different plausible values of N. The quality of the estimator method should be true for all samples.