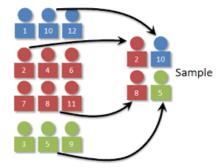# How Important is Random Sampling?
For Grades 6-8

Math Learning Goals:
- Demonstrate understanding of random sampling and nonrandom sampling.
- Discuss the importance of random sampling.
- Recall how to find mean/average, median, mode, and range and apply to data.
- Use measures of central tendency to compare datasets.

Materials Needed:
- Circle Activity Sheet
- Tree Activity Sheet
- Ruler
- Random Number Generator
- Plastic Blowup Globe (can be found online for $1.95)
-

Before the Activity:

At the beginning of class, start a discussion on random sampling. Ask the students what random sampling means, when it is used, and if they think it works. After discussing this for a little, tell the students that random sampling is a way to remove <u>bias</u> in sample selection, and tends to produce representative samples of a larger population. Ask students what happens when a personal or nonrandom selection happens instead of a random sampling.

Tree Activity: Why is Random Sampling Useful in the Real World?

The students will be finding/estimating the number of trees in a plot. The students should be given a little time to try to find a way to estimate the dots/trees from the figure below that represents an aerial photograph of tree farm. A question that should be asked is how are the dots spread over the square. To help the students, one might ask if there are spread out evenly.   You can explain that the trees are distributed approximately uniformly, meaning spread out quite evenly.

Task: Estimate the number of trees in this plot by estimating the number of dots. Your method should not involve counting a large number of dots. Counting a sample is acceptable.

Students might consider a square with sides 0.2 by 0.2 and count the dots in the square. Then multiply that total by 25 to estimate the total number of dots, since there are 25 squares of that size in the figure.
Compare the totals that different students give and discuss why they may be a bit different.

Circle Activity:

Let the students consider the 80 circles on the worksheet. We want to know: what is the average diameter for these 80 circles? Each student should follow these steps.
(1) Students can number circles in any order. Number the circles from one to 80.
(2) Take about 15 seconds and select five circles that you think best represent the size of the 80 circles.

(3) Find the average diameter for the circles in your personal sample.
(4) Find the class's average diameter. (Teacher note: the diameter is 1 cm for the small circles, 2cm for the medium-sized circles, and 3cm for the large circles).

Next, use a random digit generator to select a random sample of five circles. One way to do this is to put numbers 1-80 in a hat and draw out five. Each student should find the average diameter for the circles in his/her random sample. Then find the class's average diameter for the random sample. The sample mean diameters for the entire class can be summarized for the two selection procedures by computing the class average mean and the range of the means. One can also find the mode (most repeated average value) and median (list values in order to find one in the middle).

How do the means for the two sample selection procedures compare with the true mean diameter of 1.25 cm? Personal selection usually will tend to yield sample means that are larger than 1.25. That is, personal selection tends to be biased by systematically favoring larger circles and an overestimation of the population mean. Random selection tends to produce sample means that either underestimate the population mean or overestimate the population mean, such that the sample means cluster somewhat evenly around the population mean value. Discuss how random selection eliminates bias.

Importance and future knowledge: Imposing randomness into the sampling procedure allows us to use probability to describe the long-run variability of the sample means resulting from random sampling. The variation from repeated sampling is described through what is called the sampling distribution.

Random Sampling the Earth Activity:

Introduce a beach ball globe to the class. Ask the students if they know how much of the earth is covered by water. If there is more than one "guess", write these answers on the board. Ask the students if they can think of a way to use the beach ball globe to estimate, using random sampling, the amount of the earth that is covered by water. Discuss how much work it would be to actually measure the actual areas of the globe, since the shapes are very irregular and measuring on the surface of a sphere requires advanced math! The teacher guides the students to make a random sampling -- throw the beach ball around the class. Whoever catches it with both hands looks at where their finger lands (ie left index finger) and whether it has landed on land or water. Teacher or another student records the results. Have the students decide in advance how many times to throw it and which hand or finger to use to record the random result.
At the end, compute the probability of landing on water (number of times one lands on water divided by total number of throws).
Ask the class -- if you were to throw the beach ball another 10 times, what might happen to the probability?

TN Math Standards:
6th Grade:
Statistics and Probability:
1.) Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers.
7th Grade:
Statistics and Probability:
1.) Understand that statistics can be used to gain information about a population; generalizations about a population from a sample are valid only if the sample is representative of that population. Understand random sampling tends to produce representative samples and support valid inferences.
2.) Use data from a random sample to draw inferences about a population with an unknown

characteristic of interest. Generate multiple samples of he same size to gauge the variation in estimates or predictions.

8th Grade:

Statistics and Probability:

4.) Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables collected from the same subjects. Use relative frequencies calculated for rows or columns to describe possible association between the two variables.

The tree exercise is adapted from this textbook:

Mooney, Douglas D. and Swift, Randall J., A Course in Mathematical Modeling, Mathematical Association of America, 1999.

The circle exercise is adapted from this textbook:

Franklin, F., Kader, G., Mewborn, D., Moreno, J., Reck, R., Perry, M., and Scheaffer, R., Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report: A Pre-K-12 Curriculum Framework, American Statistical Association, 2007.

Figure 1: