

# The Forest Problem

---

Created by: Jim Spiker & Trey Cox

Chander-Gilbert Community College

[James.spiker@cgc.edu](mailto:James.spiker@cgc.edu)

## Overview of lesson

Students want to know why they would ever use a sampling method other than a simple random sample. This lesson visually illustrates the effect of using a simple random sample (SRS) vs. a stratified random sample. Students will create a SRS from a population of apple trees and use the mean of the SRS to estimate the mean yield of the trees. Students will then create a stratified random sample from the same population to again estimate the yield of the trees. The use of the stratified random sample is to control for a known source of variation in the yield of the crop, a nearby forest.

## CCSS

CCSS.Math.Content.6.SP.A.1

CCSS.Math.Content.6.SP.B.4

CCSS.Math.Content.6.SP.B.5

CCSS.Math.Content.7.SP.A.1

CCSS.Math.Content.7.SP.A.2

## Prerequisites

Students should be able to create histograms and calculate means. Students should understand the basic idea that a random sample mean estimates the population mean.

## Learning Targets

Students will be able to understand the need for using a stratified random sample to control for a known source of variation in the population.

## Time Required

Approximately 30 minutes to 1 hour.

## Materials required

Handouts of the forest problem

## Lesson Details

Provide the students with the 4x4 grid of the apple orchard. Begin by announcing the question; **Is it possible that we could estimate the mean yield of the orchard without measuring the entire output of all of the plots?** We have already discussed the idea of sampling to estimate the population so this should be rhetorical. Ask the students to design a plan for selecting 4 out of the 16 plots by any method they feel will yield a simple random sample. Discussion of the basic concept of SRS should have already taken place but the basic definition of a SRS could be reviewed here. Have the students select the plots and mark them in the grid. Now ask them what they think the effect the forest may have on the fertility of the trees near the forest versus the trees that are far away from the forest.

**Is it reasonable to assume the apple trees near the forest may be affected in some way by the presence of the forest?**

**Do they think the trees near the forest may be more or less likely to produce more fruit?**

Any reasonable opinion is worth exploring here. Now ask them **how they might be able to select their simple random sample by stratification. What if we break the block of 16 plots up into four horizontal rows and select one plot at random from each row? Could this control for the effect of the forest? Or what if we could break the plot up into four vertical columns and select one from each column?** This will be the best method as the students should discover for themselves when they actually measure the average yield of their samples and construct a sampling distribution.

Now have the students turn to the next page of the handout. This gives the number of bushels of apples the plot of land produced. Have the students find the average of their four plots and come mark their average as a dot on the dot plot for the entire class. When you have done this for each of the three different samples you should **then have the class discuss the amount of variability they see in each sampling distribution.** You should allow them to find the mean of the sample means for each type of distribution. All three means should be near 70, however, the first and the second histograms will have a large amount of variation. You could have the students calculate the MAD for each distribution. You will find the MAD for the SRS and the stratified random sample by rows will have a much greater MAD and hence much more variability than the stratified sample based on columns. Use this time **to discuss why pulling one from each column controls for the variation by forcing the sample to contain one plot from each distance from the forest**, while the SRS and the stratified by rows could allow a sample to be entirely near the forest or entirely far away from the forest by random chance.

## Student Handout

1. Assign the numbers 1-16 to the plots below. Using your TI-83/84 type `math-prb-randint(1,16,4)` to pick a simple random sample (SRS) of 4 plots. Place an X in the 4 plots that you choose.




2. Assign the numbers 1-4 to each horizontal row. Now, randomly choose one plot from each horizontal row by `Math-prb-randint(1,4,4)`. This is called a stratified random sample.




3. Now assign the numbers 1-4 to each vertical column and pick a random sample of one plot from each column. Randomly choose one plot from each vertical column. This is also a stratified random sample.




4. Now, it is for the harvest! The numbers below are the yield for each plot.

4	29	94	150
7	31	98	153
6	27	92	148
5	32	97	147

