



# Stratified Sampling

## Student Activity

Name \_\_\_\_\_

Class \_\_\_\_\_

Open the TI-Nspire document *Stratified\_Sampling.tns*.

In this activity, you will determine which of three different sampling methods—a simple random selection design and two stratified selection designs—would be most beneficial in selecting a survey sample within a given context.



**Move to page 1.2.**

Press **ctrl** and **ctrl** to navigate through the lesson.

Pages 1.2 and 1.3 contain directions to “seed” your handheld for the activity.

The Happy Sundae Ice Cream Company is looking to expand its ice cream truck business into new neighborhoods. From previous business experience, the company knows how many children need to live in a neighborhood for an ice cream truck to be profitable.

The neighborhood being considered is bounded by the downtown on the south and a freeway on the east and consists of 36 approximately square city blocks. Sean, Happy Sundae’s Director of Marketing, wants to survey six of the city blocks to get an estimate of the total number of children living in the neighborhood.

**Move to page 1.4.**

1. The first sampling method Sean decides to use is a simple random sample (SRS). The left work area displays the 36 city blocks in which Sean is interested. Use the right arrow once to randomly select six city blocks to survey (you are actually *simulating* what would happen if these surveys were really conducted).

Note: The left arrow will erase all the samples and allow you to begin again.

- A number is displayed in each "block", representing the simulated number of children in the city block surveyed.
  - The top dotplot in the right work area of the screen displays the number of children found in each of the city blocks surveyed.
  - The lower dotplot displays the mean of the sample.
- a. The vertical line on the top dotplot represents the mean number of children in the sample of six city blocks surveyed. This value for the mean also appears as a dot in the lower half of the screen. What would a mean of 40 represent in the context of the entire neighborhood and the ice cream truck business?



# Stratified Sampling

## Student Activity

Name \_\_\_\_\_

Class \_\_\_\_\_

- b. Click on the right arrow to create another simulation of a simple random sampling for 6 city blocks. Write down any observations you have regarding the number of children in the neighborhood for this sample.
- c. Click the arrow to create 23 more simulated samples (of 6 city blocks each) for a total of 25 simulations, each time observing the number of children in each city block sampled. Fill in the table below with the results from each simulated sample. The number of children in each city block is fixed, so you only fill in a cell once. What pattern(s) do you see in the number of children? Explain your reasoning.


DOWNTOWN

FREEWAY

- d. What is the mean number of children per city block? Describe the number of children possibly living in the entire 36-block neighborhood according to this strategy.
2. In thinking about the neighborhood, Sean suspects that there might be fewer children in the blocks close to the freeway and that a simple random sample might not be the best sampling method to use. Not having a block near the freeway as part of the sample could potentially overestimate the mean number of children in the neighborhood. He decides to redesign his sampling to account for this factor. Stratified random sampling involves dividing the population into similar sub-populations or strata before sampling. The strata need to be as homogeneous (similar) as possible and not overlap.
- a. If Sean suspects that there are fewer children closer to the freeway, describe how he might create strata to reflect this and explain your reasoning.



**Move to page 1.5.**

- b. This page shows a simulation of stratified random sampling using the freeway as the factor. Six city blocks need to be selected. How would you suggest selecting them using vertical strata?
  
  - c. Click on the arrow to see the number of children per block in a stratified random sample of city blocks surveyed (freeway as the stratifying factor). Observe the plot and the dotplot on the right, and write down any observations you have about the number of children.
  
  - d. Click the arrow 24 more times for a total of 25 simulations, each time observing the mean number of children in the plot and the dotplot on the right. The vertical line in the top panel represents the mean of the sample. What is the mean number of children per city block? Describe the possible number of children in the entire neighborhood if using this stratified sampling strategy. Compare this to the answer you found for simple random sampling.
  
  - e. Explain the variation in the plot of means.
3. Sean remembers an article he read in last Sunday's newspaper that suggested there are fewer children in neighborhoods farther from downtown. He decides to redesign his sampling one more time based on this information.
- a. The strata need to be as homogeneous as possible and not overlap. If Sean suspects that there are more children in neighborhoods closer to downtown and fewer children farther from downtown, describe how to create six strata to reflect this and how to choose a sample from the strata.



**Move to page 1.6.**

- b. Click on the arrow 25 times to see a simulated sampling distribution of the number of children in a stratified random sample (downtown as the stratifying factor) of city blocks surveyed. Observe the variation in the mean number of children per city block for this stratified sampling strategy. How does the mean number of children compare to the means you found earlier? What would this suggest about the total number of children in the 36-block neighborhood?

**Move to page 1.7.**

4. This page shows boxplots of the sample means for all three of the sampling designs used—simple random sample, stratified vertically (parallel to the freeway), and stratified horizontally (parallel to downtown).
  - a. Compare the centers of the three boxplots. Are they similar, different? How many total children does each represent? Explain your answers.
  - b. Compare the variation in the three boxplots.
  - c. Which factor, the freeway or downtown, seems to be relevant to the sample selection? Explain.
5. The actual mean is 34.89 children per block. How many total children does this represent for the entire neighborhood? Do each of the sampling designs reflect this total? Explain.
6. The goal of the survey was to estimate how many children are in the neighborhood. Sean will only get to select one sample on which to base his decision regarding the new ice cream route. Which method—simple random sample, stratified horizontally, or stratified vertically—should the Director of Marketing use to most effectively eliminate the influence of a factor? Take into consideration the center and the spread of the simulated sampling distributions as you explain your reasoning.