Sampling Methods – Stratified

The goal of this activity is to use simulations to explore the sampling methods of *stratified random samples*.

Suppose you were again using a sample to estimate the population mean of the areas of rectangles shown on the activity sheet "Random Rectangles," and this time you wanted to make sure that your sample contained rectangles with both small and large widths.

Tables 1 and 2 show the set of 100 rectangles divided into two groups (strata) according to the widths of the rectangles. Each number corresponds to the rectangle on the activity sheet "Random Rectangles." Note that in this division of rectangles, width indicates a rectangle's horizontal dimension. One stratum contains 59 rectangles with widths less than 3 (Table 1), and the other stratum contains 41 rectangles with widths greater than or equal to 3 (Table 2).

Table 1. Stratum of Rectangles with Widths Less than 3

1	2	3	4	5	6	8	9	10	11	12	14	16	18	19
20	21	22	24	27	28	30	32	34	358	36	40	43	44	45
46	48	49	50	51	53	56	62	64	67	68	69	71	73	74
75	77	78	79	80	82	83	84	87	88	89	91	95	100	

Table 2. Stratum of Rectangles with Widths Greater than or Equal to 3

7	13	15	17	23	25	26	29	31	33	37	38	39	41	42
47	52	54	55	57	58	59	60	61	63	65	66	70	72	76
81	85	86	90	92	93	94	96	97	98	99				

Stratified random sampling is a process that involves randomly selecting samples from groups like these, each of which is considered a homogeneous stratum of the population.

Use the rectangles in Tables 1 and 2 to explore the process and results of stratified random sampling:

1. Using some type of random number generator, randomly select five rectangles from each table (stratum), and then compute the mean area of each stratum of five rectangles. The sample mean for the combined strata is found by using the population proportion, as follows:

$$\frac{59}{100} \cdot (\text{mean of stratum from Table 1}) + \frac{41}{100} \cdot (\text{mean of stratum from Table 2})$$

Stratified Rectangles

Stratum 1	Area	Stratum 2	Area
1. _____	_____	1. _____	_____
2. _____	_____	2. _____	_____
3. _____	_____	3. _____	_____
4. _____	_____	4. _____	_____
5. _____	_____	5. _____	_____

Sample mean area of the stratified rectangles _____

Source: *Navigating through Data Analysis in Grades 9–12*, NCTM

Sampling Methods – Clustered

The goal of this activity is to use simulations to explore the sampling methods of *clustered samples.*

Sometimes a population of interest has so many members or the members are so dispersed that the cost of taking a simple random sample is too high. In such instances, researchers may use cluster sampling, a random selection process by which clusters of individuals are identified in the population, and the individuals in the clusters are studied.

Table 3 shows the 100 rectangles for the activity sheet "Random Rectangles" divided into 20 clusters. Each cluster is "like" every other cluster in that it contains five rectangles that are relatively close (clustered) together on the sheet. Each also contains as much variability as is possible in its particular "neighborhood." Each cluster is named by a Roman numeral.

Table 3. Clusters of Rectangles Listed by the Numbers on "Random Rectangles"

I	II	III	IV	V	VI	VII	VIII	IX	X
1	3	8	7	19	18	25	40	32	39
2	4	13	11	26	20	34	41	33	47
9	5	14	12	27	21	35	42	38	48
16	6	15	22	28	30	36	43	45	49
17	10	24	23	29	31	37	44	46	50
XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
51	55	58	68	64	66	79	91	88	83
52	56	59	69	72	75	80	92	89	84
53	57	60	70	73	76	81	93	96	90
54	62	65	71	74	77	86	94	97	99
61	63	67	78	82	83	87	95	98	100

Use the clusters of rectangles in Table 3 to explore the process and results of cluster sampling:

2. Using a random number generator, select two of the 20 clusters. Find the mean area of your sample of 10 rectangles from the two clusters.

Cluster Rectangles

Cluster _____	Area	Cluster _____	Area
1. _____	_____	1. _____	_____
2. _____	_____	2. _____	_____
3. _____	_____	3. _____	_____
4. _____	_____	4. _____	_____
5. _____	_____	5. _____	_____

Sample mean area of the cluster rectangles _____

Source: *Navigating through Data Analysis in Grades 9–12*, NCTM

Sampling Methods – Systematic

The goal of this activity is to use simulations to explore the sampling methods of *systematic samples.*

In simple random sampling, stratified random sampling, and even cluster sampling, it is necessary that a list of individuals in the population being studied exists (to do the sampling properly). A sampling technique that does not require this is systematic sampling.

A systematic sample is obtained by selecting every *k*th individual from the population. The first individual selected corresponds to a random number between 1 and *k*.

Please refer to the activity sheet "Random Rectangles" to complete the following:

3. Select a systematic sample of size 10 from the population of 100 rectangles using the following procedure.
   a. Find the value of *k* by dividing the population size by the desired sample size.
   b. Using a random number generator, select a number between 1 and *k*. Call it *j*.
   c. Using the number *j* found in (b), add successive multiples of *k* to obtain the 10 numbers. Record this sequence of numbers below.
   d. Use the numbers to locate the 10 rectangles that have the corresponding numbers on the sheet, then compute and record the sample mean area of the 10 rectangles.

Systematic Rectangles

Value of *j* = ___	Area		Area
1. _____	_____	6. _____	_____
2. _____	_____	7. _____	_____
3. _____	_____	8. _____	_____
4. _____	_____	9. _____	_____
5. _____	_____	10. _____	_____

Sample mean area of the systematic rectangles _____

Source: *Navigating through Data Analysis in Grades 9–12*, NCTM

4. For each of the sampling methods above, record the results from each member of your group.

a. Stratified Sampling:

Mean of Stratum 1	Mean of Stratum 2	Mean of Combined Strata
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

b. Cluster Sampling:

First Cluster	Second Cluster	Mean area of Two Clusters
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

c. Systematic Sampling:

Areas of rectangles starting with rectangle *j*	Mean area of Sequence
_____	_____
_____	_____
_____	_____
_____	_____