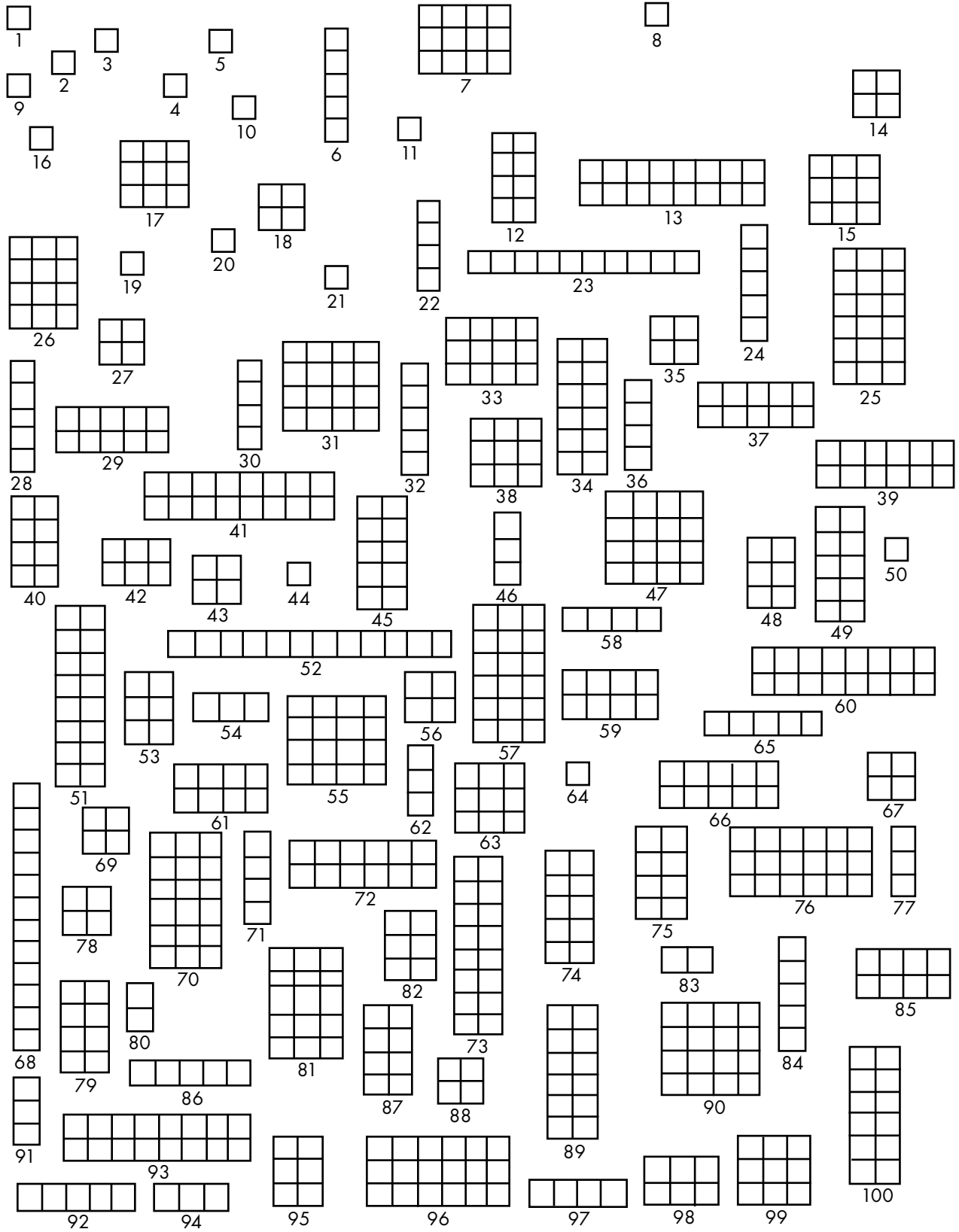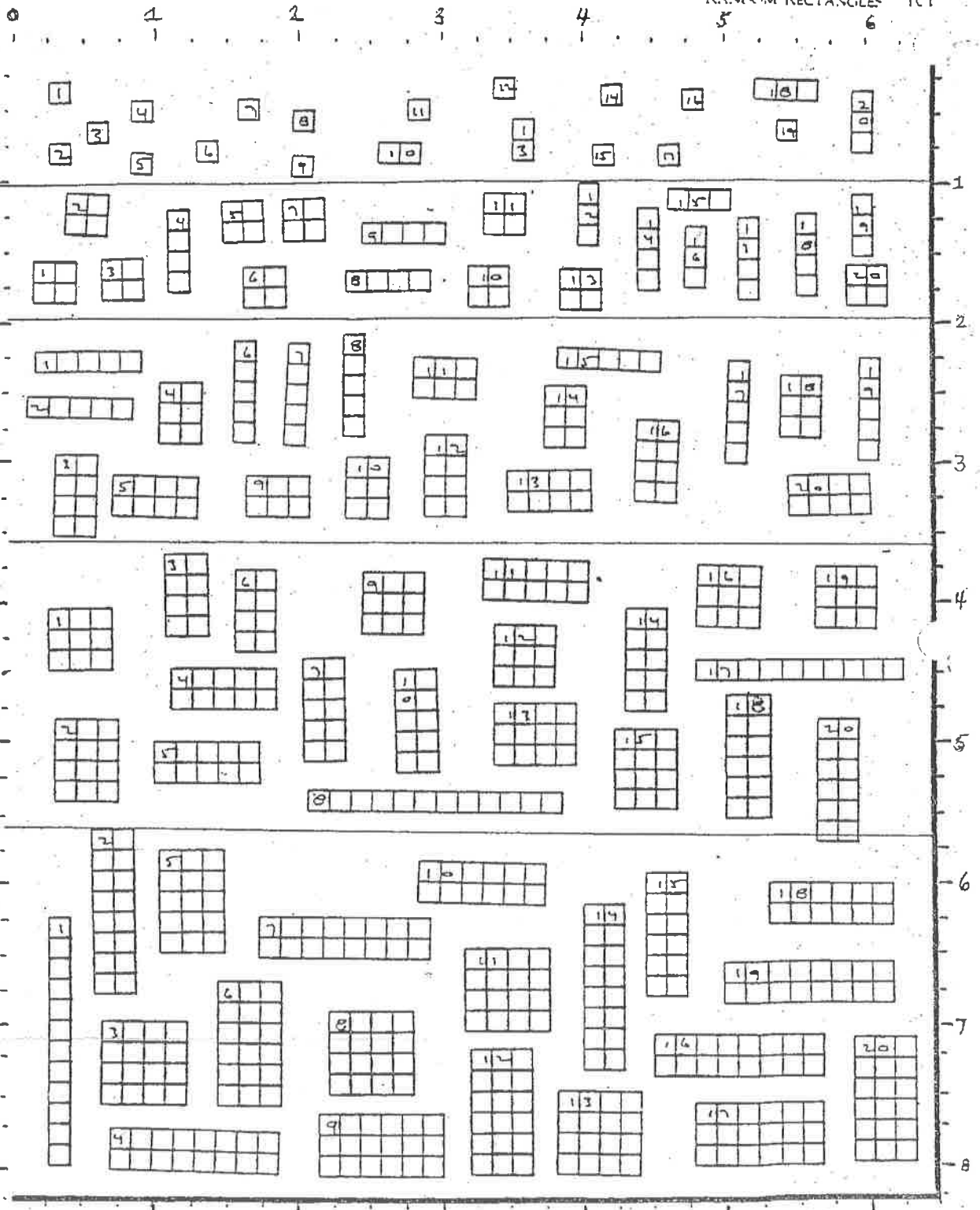# 100 Random Rectangles

Random Rectangles Activity

Objective: Understand how subjective samples will be compared to random samples in terms of sampling bias. The goal is to learn why randomization is an important part of data collection. Additional goals include generating data sets that follow an approximately normal distribution and seeing how sample size affects our confidence in making predictions.

Activity:

(1) Do not look at the back of this paper until the instructor gives the signal to begin. Then look at the sheet for a few seconds and write down your guess as to the average area of the rectangles on the sheet. (Each small square is one square unit.)

Guess of Average Area: _____

(2) Now select five rectangles that, in your judgment, are representative of the rectangles on this page. Write down the area for each of the five. Compute the average of the five areas, and compare it to your guess. Are the averages for (1) and (2) close?

Rectangle #					
Area of Rectangle					

Average Area: _____

(3) Now that you have answered questions (1) and (2), write your averages on the board. We'll use this whole-class data later.

(4) Now use a random number generator to select 5 distinct random rectangles between 1 and 100. Then find the average area of these 5 rectangles.

Random Rectangle #					
Area of Rectangle					

Average Area: _____

(5) Repeat step (4) several more times, and record your average area each time. When you are finished, record your data on the board for use later.

(6) Repeat step (4) using a sample of 10 distinct random rectangles and compute the average area. Do it a couple times if you can, and record the data on the board for use later on.

Analysis:
Using the whole-class data that is on the board, calculate the means, standard deviations, and five-number summaries for these distributions. (You may use a calculator). Here are some questions to consider:

  a. How do the centers and spreads of the various distributions compare?
  b. Which method of sampling (subjective or random) do you think is doing a better job? Why?
  c. How does the amount of spread in the 10-rectangle-based sampling distribution compare with the 5-rectangle-based sampling distribution? Which distribution gives a more trustworthy estimate of the true population mean?