Learning Goal: For the distribution of a quantitative variable, describe the overall pattern (shape, center, and spread) and striking deviations from the pattern.

Specific Learning Objectives:
- Develop a way to describe and distinguish graphs of a quantitative variable.
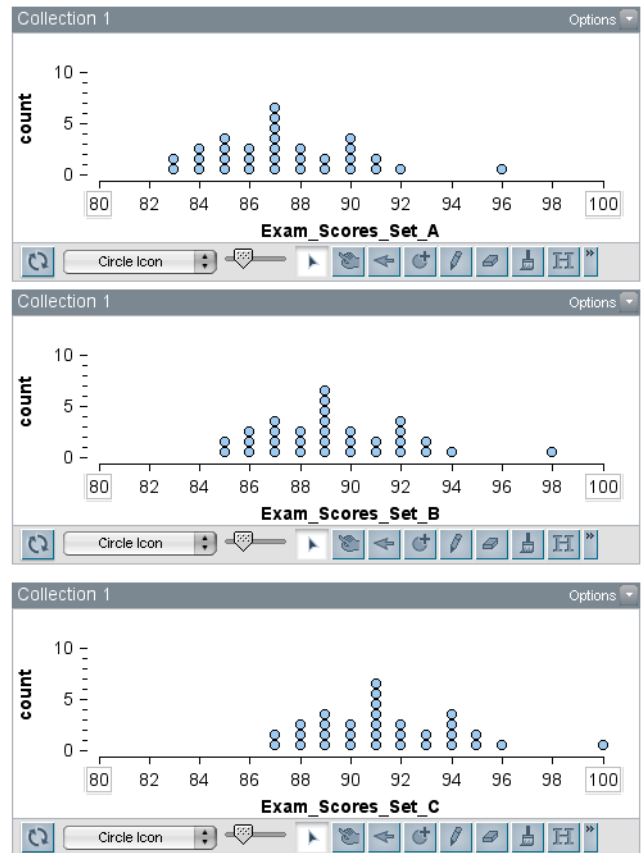- Identify reasonable explanations for what might explain the differences seen in different data sets.

Introduction: Statisticians collect data in order to study and analyze groups. They focus on the group's data in aggregate, instead of focusing on individuals. For a statistician creating and comparing graphs is often the first step in analyzing data.

In this activity you will compare and contrast the graphs of hypothetical sets of exam scores. In other words, these graphs show made-up data. These hypothetical data sets have been constructed to help you begin to "see" like a statistician. By comparing these data sets, you will begin to develop an informal understanding of the key features of a graph that statisticians use to describe data. We will develop these ideas more formally in future lessons.

During this activity do your best to describe what you see. Jot down notes to capture your thinking as you go. You can use your notes during our class discussion of this activity. This activity does not require you to remember anything or to apply previous knowledge.

1) Compare and contrast the 3 graphs shown at the right.

Exam_Scores_Set_A

Exam_Scores_Set_B

Exam_Scores_Set_C

a) How are the graphs similar? How are they different? What is the most distinctive feature that distinguishes these three graphs from each other?

b) For each graph, pick a single exam score to summarize the overall performance of the students. In other words, summarize each set of data with one number.
   Set A _____ Set B _____ Set C _____

c) The average score for Set A is 87.4. What do you think the average score is for Set B and for Set C? (See if you can answer this without doing any calculations.)

d) Which, if any, of the statements below is a reasonable explanation for the differences in the graphs? Why?

   • The graphs represent different classes. Different groups of students will obviously perform differently on an exam.

   • The graphs represent a single class after the teacher adjusted the grades. The teacher realized that some of the exam questions were not well written, so she adjusted the grades by adding points to the original exam scores.

   • There are no differences in the graphs. They look the same.

2) Compare and contrast the 3 graphs shown at the right.

a) How are the graphs similar? How are they different? What is the most distinctive feature that distinguishes these three graphs from each other?
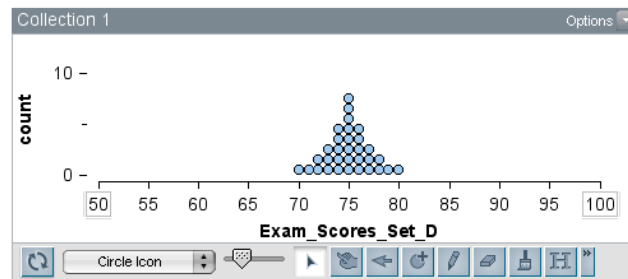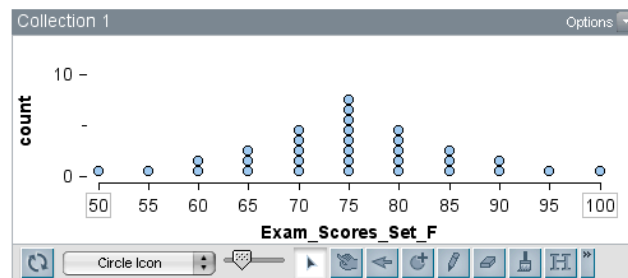

b) *Thinking about center:* The average exam score is the same for each set of data (D, E and F). What do you think the average is?
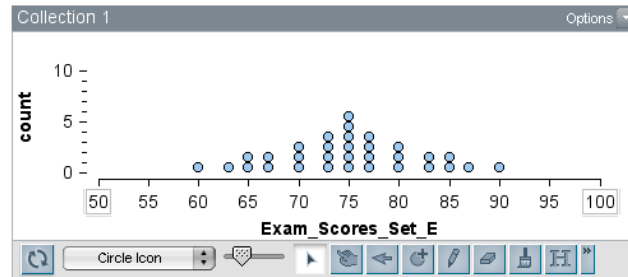
c) *Thinking about spread:* Spread is a description of the variability we see in the data.
For exam set D the scores vary from a low of _____ to a high of _____.

For exam set E the scores vary from a low of _____ to a high of _____.

For exam set F the scores vary from a low of _____ to a high of _____.

d) Which of the statements below, if any, is a reasonable explanation for the differences in the graphs? Why?

- The graphs represent a single class after the teacher adjusted the grades. The teacher realized that some of the exam questions were not well written, so she adjusted the grades by adding points to the original exam scores.

- The graphs represent 3 different classes with teachers that have different grading standards. One is an easy grader. One is a really hard grader.

- The differences could be explained by whether the teachers allowed the students to work together on the exam and how much time the students were given to finish the exam.

3) Compare and contrast the 3 graphs shown at the right.

a) In each of these 3 graphs, the average score is the same. The average score is 75. In what other ways are the graphs similar? How are they different? What is the most distinctive feature that distinguishes these three graphs from each other?
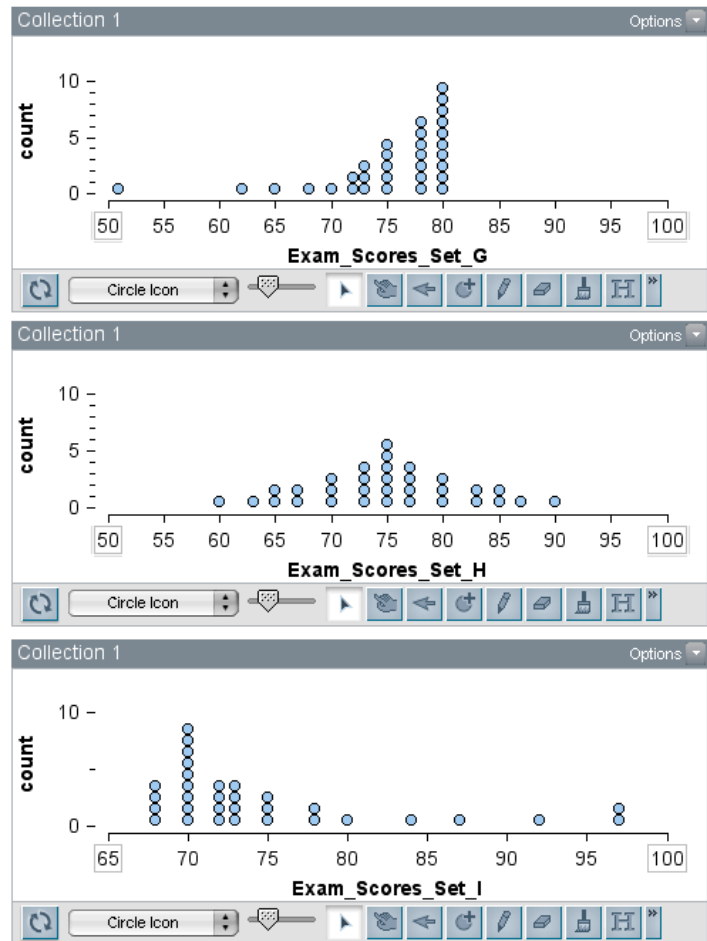

b) What might explain the differences in these graphs?

c) *Thinking about shape:* How would you describe the shapes (symmetric, skewed left, skewed right) of these three distributions of exam scores?

4) For each of the following dotplots, draw a smooth curve outlining the distribution, and then describe the shape of the distribution using course vocabulary (See the definitions of symmetric, skewed left, skewed right.)

Bicep girth for 247 men who were exercising several times a week.
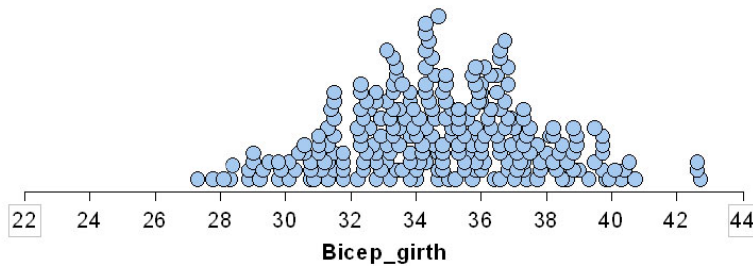

Waist girth for 260 women who were exercising several times a week.
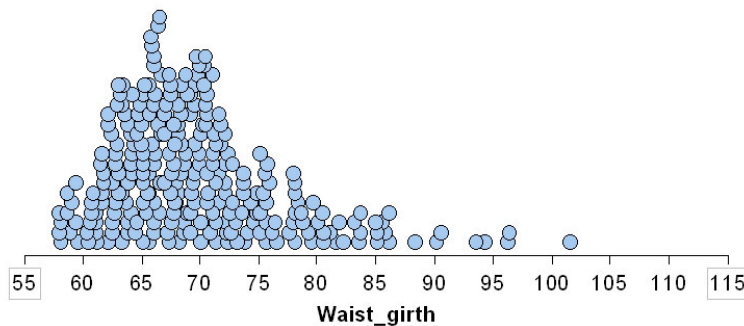
5) Suppose that almost everyone does well on the first exam with a few people (who did not study) performing poorly. What is the shape of the distribution of exam scores?

6) Here are city mileage ratings for 44 hatchback cars.

   Thinking about deviations from the pattern: Are there any unusually large or unusually small mpg rates? If so, what are the unusual values?
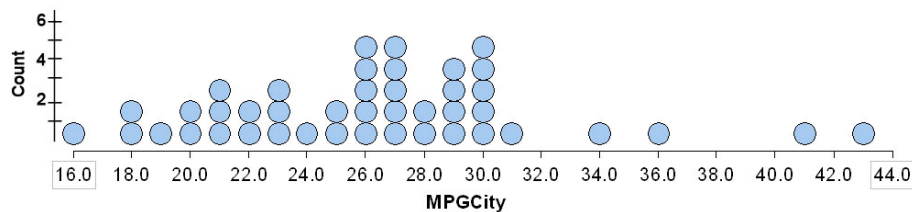
7) Here are dot plots of the sugar content (grams per serving) for some adult cereals and child cereals. Compare the two distributions by comparing shapes, estimates of center, and spreads.