**Learning Goals**

- Create and interpret different graphs of a quantitative variable.
- Summarize and describe the distribution of a quantitative variable in context. Describe the overall pattern (shape, center and spread) and striking deviations from the pattern.

**Specific Learning Objectives**

- Use quartiles to quantify variability relative to the median.
- Create and interpret boxplots, relate boxplots to histograms and dotplots.

1) Recall that the median is one way to summarize a distribution with a single number. Half of the data lies above the median and half lies below it. The median can be an actual data value or it can be a halfway mark between two values.

   a) Find the median for this set of 23 numbers:

   5 5 6 6 6 7 9 9 11 13 13 14 15 15 15 15 16 17 19 20 21 22 22

Statisticians measure spread relative to the median by marking the quartiles. Quartiles divide the data into four groups, with 25% of the data in each group (or as close to 25% as possible). The first quartile mark (Q1) has 25% of the data below it (or as close to that as possible). The second quartile mark (Q2, the median) has 50% of the data below it (or close to that as possible). The third quartile mark (Q3) has 75% of the data below it (or close to that as possible). You can think of Q1 as the median of the lower half of the data. Similarly, Q3 is the median of the upper half of the data.

   b) The 5-number summary is the minimum, Q1, Q2, Q3, maximum. Find the 5-number summary for the data set above.

   c) The 5-number summary can be used to make a boxplot. Here is the boxplot for this set of numbers. Study it and see if you can explain how to make a boxplot based on your work in part b.
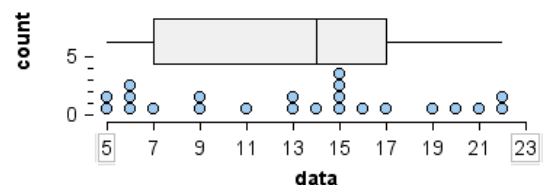
2) Draw a number line and make a boxplot to represent a data set with this 5-number summary: 3, 7, 9, 15, 20. (Make sure your number line has a constant scale. For example, you could mark off the number line in increments of 2.)

3) Here is a data set with its dot plot. We also made a histogram and a boxplot from the same data.

5  6  6  7  8  8  8  9  10  14  15  17  18  20  21
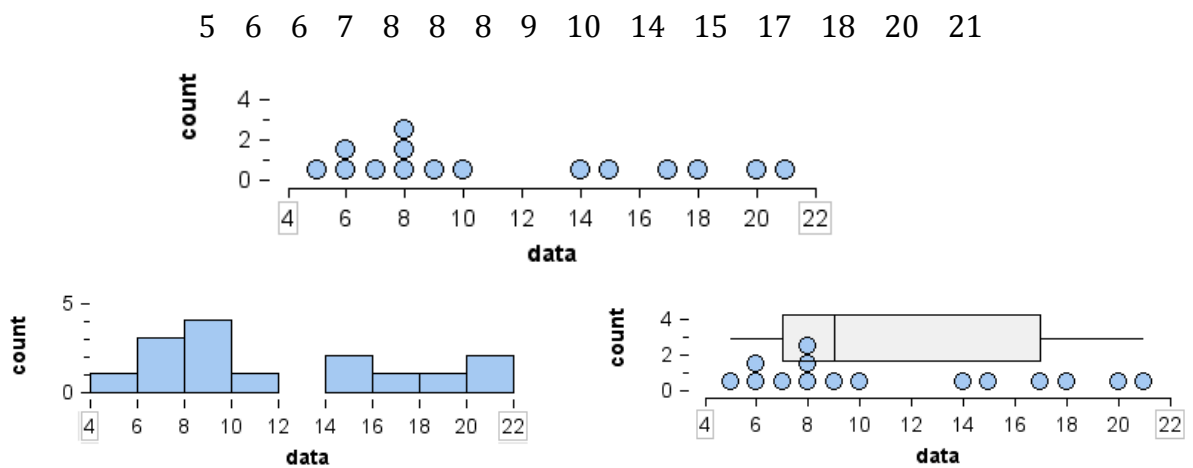


a) Use the list of data below to show how the data is grouped to make the bins in the histogram. In other words, circle the groups of data for each bin.

5  6  6  7  8  8  8  9  10  14  15  17  18  20  21

b) Use the list of data below to show how the data is grouped to make the boxplot. In other words, circle the groups of data for each quartile.

5  6  6  7  8  8  8  9  10  14  15  17  18  20  21

c) The graph that divides the data into four groups of equal counts is (circle one: histogram, boxplot). The graph that divides the data into bins of equal widths is (circle one: histogram, boxplot).

d) What is the 5-number summary for this data set? (Hint: you should be able to look at the boxplot to determine this.)

4) Match the histograms to the boxplots.
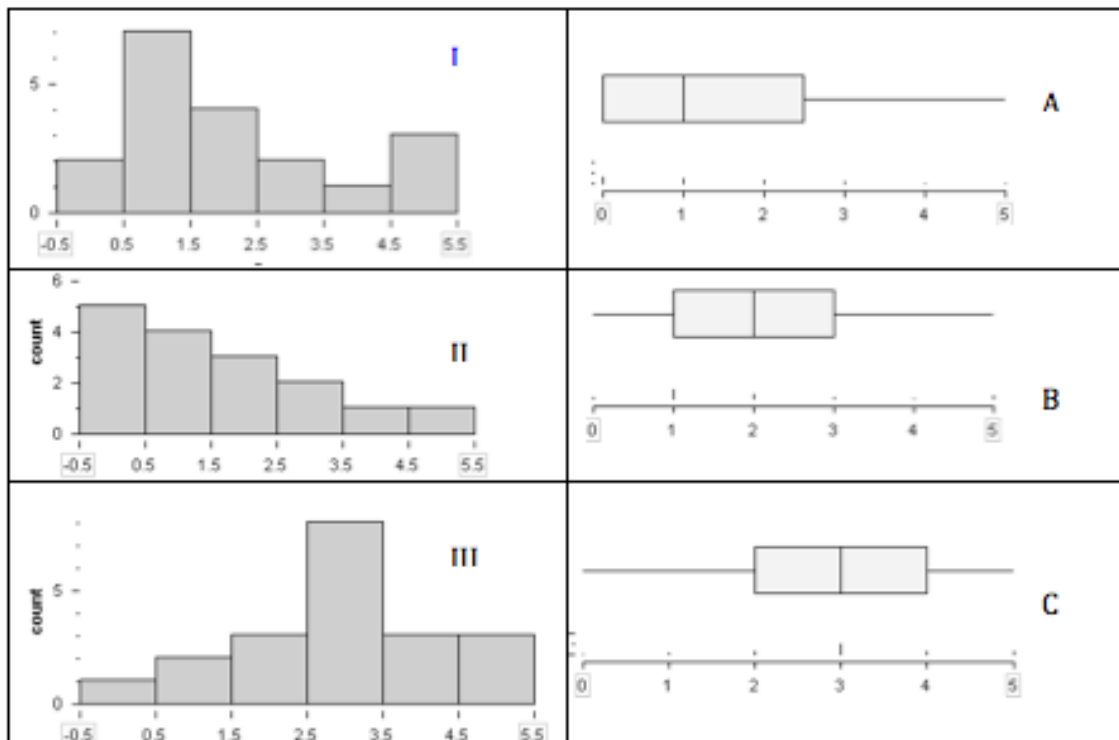


5) A histogram and a boxplot are two ways to summarize a distribution. Distributions with similar patterns, but different values, can produce the same histogram or the same boxplot. In other words, we lose information about the actual data by summarizing it into a graph. To illustrate this idea, make up a data set with 11 data points (n=11) that fits this pair of graphs. Then find a different data set (n=11) that also fits these two graphs.
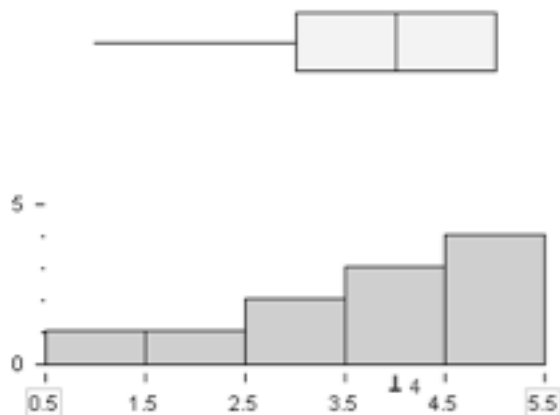
6) Make up a data set (n=11) that fits this boxplot and has the most amount of spread relative to the median as possible. Make a dot plot below the boxplot on the left to show your data. Then make up a second data set (n=11) that fits this boxplot and has the least amount of spread relative to the median as possible. Make a dot plot below the boxplot on the right to show your data.
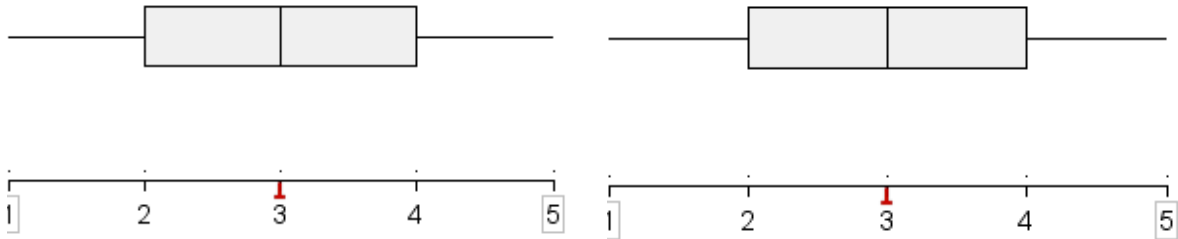
7) Make up a data set with 11 numbers that matches this boxplot. Make a histogram of your data.
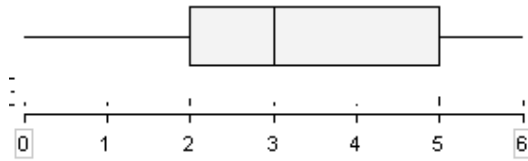
8) Draw a boxplot for the data shown in the dot plot.