Name:_____ Date:_____

**AP Statistics: Assignment #1.1 Boxplots, Exploring Data**

1. What are boxplots? What are they used for? Explain:

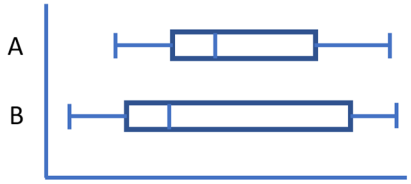
2. How do you find the mean, median, mode, Q1, Q3, max, and min when looking at a boxplot used at describing a distribution? Explain:

3. What is an outlier? How do you classify a data point as an outlier?

4. What is an IQR? What is it used for? Explain:

5. When comparing two boxplots or distributions, what information should be provided?

6. Given the two boxplots below, describe and compare the two distributions:




7. Suppose the box plot above represents the distribution of math scores from two different classes. What information can you use to claim that class "A" did better?

8. In contrast, using the box plots from Q6, what information can you use to claim that class "B" did better?

9. A group of students from AP Statistics were timed when asked to complete the last question from the Gauss math contest. Use the stem leaf plot to answer the questions below:

| | |
|---|---|
| 0 | 9 |
| 1 | 1 1 2 2 3 4 |
| 1 | 5 5 7 9 9 |
| 2 | 1 1 3 3 |
| 2 | 6 8 9 |
| 3 | 1 |
| 3 | 7 |
| 4 | 1 |
| 5 | 9 |

a) Find the 5 number summary: Min, Q1, Median, Q3, and Max

b) Find the IQR and identify any outliers

c) Create a boxplot and then explain how the mean compares to the median

d) Use your Ti83 to calculate and then interpret the mean.

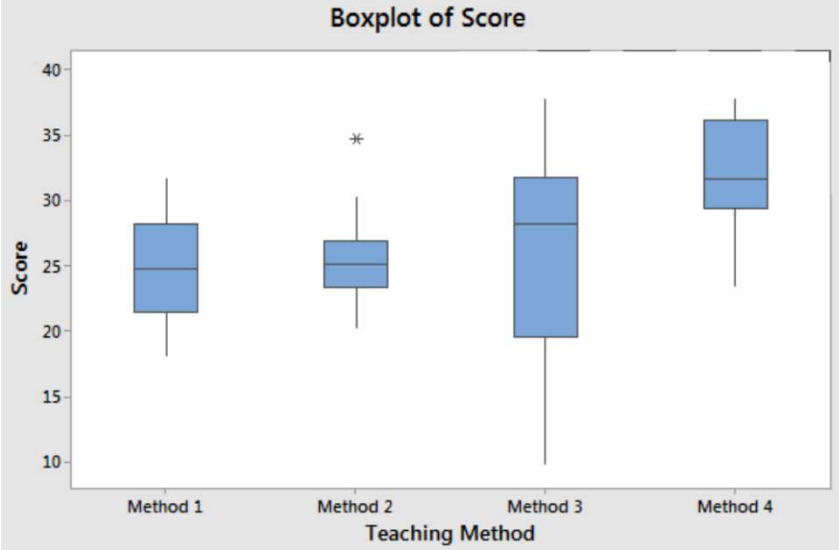
e) Use your Ti83 to calculate and then interpret the median

f) Which measure of center would be the most appropriate summary of the center of this distribution? Explain why:

10. The number of pages of Mr. Cheong's favorite books are noted below:
240 , 350, 310, 346, 320, 286, 336, 366, 190, 354, 318, 376, and 330.
    a) Use the data above to construct a boxplot. Describe the center and spread using the five number summary:

    b) Calculate and interpret the mean and standard deviation for these data:

11. The following graph shows the distribution of scores from four classes that utilized different teaching methods. Each class had a class size of 50 students and were randomly assigned. Student scores ranged from 0 to 40. Use the graph to answer the following questions:




a) Describe the shape, center, and spread of the distribution of scores for each teaching method

b) What do the data show, how would you rank the efficiency of teaching method? Justify your answer.

12. For each of the following statements, indicate whether if it is true or false.
    a. If a distribution is skewed left, then the median is larger than the mean _____

    b. If a distribution is skewed right, then the median is larger than mean _____

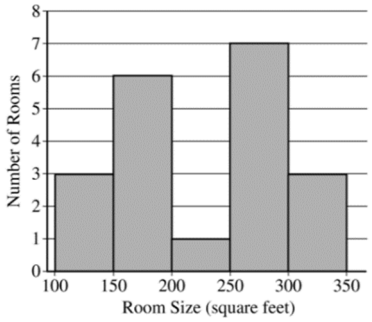
    c. If the mean is larger than the median, then the distribution is skewed right _____

    d. If the mean is smaller than the median, the the distribution is skewed left _____

    e. If a density curve is symmetric, then the mean is equal to the median _____

13. If a set of data shows a distribution that is skewed right, would the “mean” or “median” be a better measure of central tendency? Explain:

14. The sizes, in square feet, of the 20 rooms in a student residence hall at a certain university are summarized in the following histogram. The summary statistics for the sizes are also given in the table below:




| Mean | Standard Deviation | Min | Q1 | Median | Q3 | Max |
|---|---|---|---|---|---|---|
| 231.4 | 68.12 | 134 | 174 | 253.5 | 292 | 315 |

a) Using the histogram, write a few sentences describing the distribution of room size in the residence hall

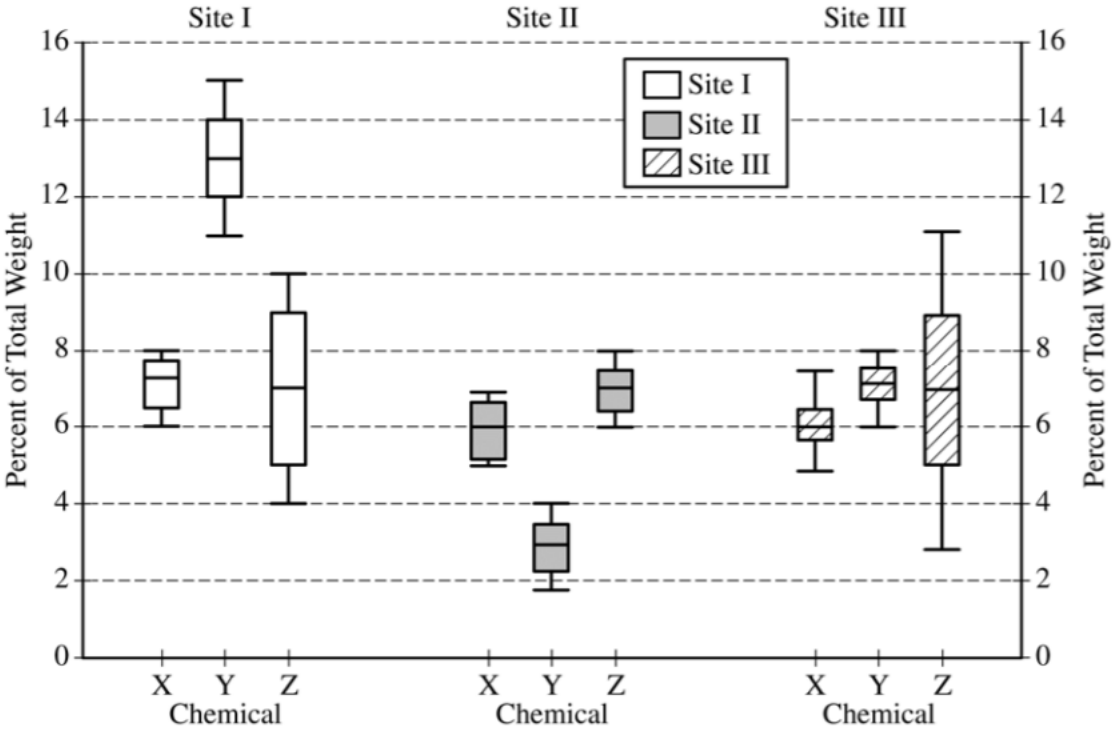
b) Determine whether there are potential outliers in the data. Use the grid below to draw a boxplot of the room size:

c) What characteristic of the shape of the distribution of room size is apparent from the histogram but not from the boxplot?

Question7

The chemicals in clay used to make pottery can differ depending on the geographical region where the clay originated. Sometimes, archaeologists use a chemical analysis of clay to help identify where a piece of pottery originated. Such an analysis measures the amount of a chemical in the clay as a percent of the total weight of the piece of pottery. The boxplots below summarize analyses done for three chemicals—X, Y, and Z—on pieces of pottery that originated at one of three sites: I, II, or III.

(cont. on next page)




(a) For chemical Z, describe how the percents found in the pieces of pottery are similar and how they differ among the three sites.

(b) Consider a piece of pottery known to have originated at one of the three sites, but the actual site is not known.

z

(i) Suppose an analysis of the clay reveals that the sum of the percents of the three chemicals X, Y, and Z is 20.5%. Based on the boxplots, which site—I, II, or III—is the most likely site where the piece of pottery originated? Justify your choice.

(ii) Suppose only one chemical could be analyzed in the piece of pottery. Which chemical—X, Y, or Z—would be the most useful in identifying the site where the piece of pottery originated? Justify your choice.