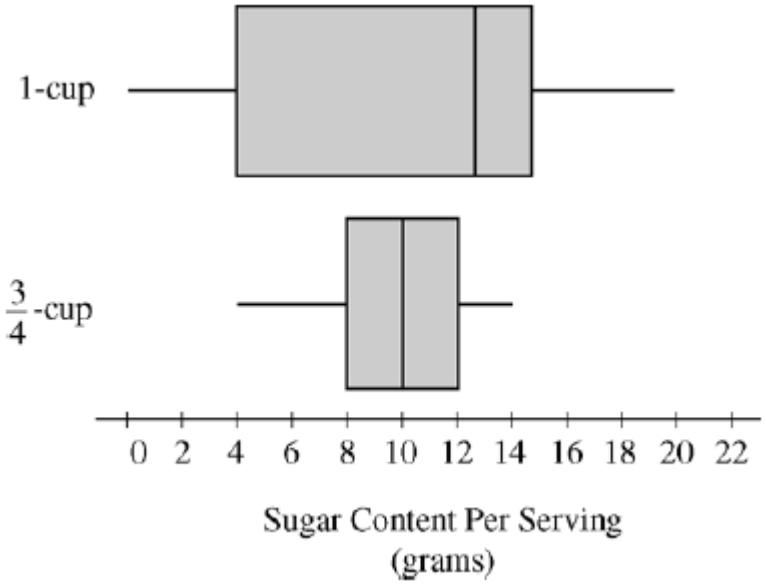
Name: _____ Date: _____ Bell: _____

AP Stats POW #10

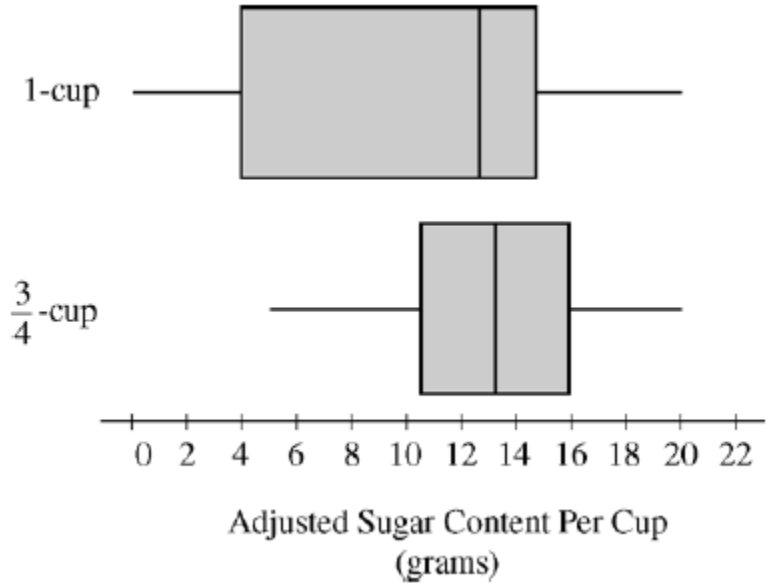
Answer each question as thoroughly as possible.

1. To determine the amount of sugar in a typical serving of breakfast cereal, a student randomly selected 60 boxes of different types of cereal from the shelves of a large grocery store.
   The student noticed that the side panels of some of the cereal boxes showed sugar content based on one-cup servings, while others showed sugar content based on three-quarter cup servings. Many of the cereal boxes with side panels that showed three-quarter cup servings were ones that appealed to young children, and the student wondered whether there might be some difference in the sugar content of the cereals that showed different size servings on their side panels. To investigate the question, the data were separated into two groups. One group consisted of 29 cereals that showed one-cup serving sizes; the other group consisted of 31 cereals that showed a three-quarter cup serving sizes. The boxplot shown below display sugar content (in grams) per serving of the cereal for each of the two serving sizes.



   a. Write a few sentences to compare the distributions of sugar content per serving for the two serving sizes of cereals.

After analyzing the boxplot on the preceding page, the student decided that instead of a comparison of sugar content per recommended serving, it might be more appropriate to compare sugar content for equal size servings. To compare the amount of sugar in serving sizes of one cup each, the amount of sugar in each cereal showing three-quarter cup servings on their side panels was multiplied by $\frac{4}{3}$. The bottom boxplot shown below displays sugar content (in grams) per cup for those cereals that showed a serving size of three-quarter cup on their side panels.



b. What new information about sugar content do the boxplot above provide?

c. Based on the boxplots shown above on this page, how would you expect the mean amount of sugar per cup to compare for the different recommended serving sizes? Explain.

2. The depth from the surface of Earth to a refracting layer beneath the surface can be estimated using methods developed by seismologists. One method is based on the time required for vibrations to travel from a distant explosion to a receiving point. The depth measurement (M) is the sum of the true depth (D) and the random measurement error (E). That is, M = D + E. The measurement error (E) is assumed to be normally distributed with mean 0 feet and standard deviation 1.5 feet.
   a. If the true depth at a certain point is 2 feet, what is the probability that the depth measurement will be negative?

   b. Suppose three independent depth measurements are taken at the point where the true depth is 2 feet. What is the probability that <u>at least one</u> of these measurements will be negative?

   c. What is the probability that the <u>mean</u> of the three independent depth measurements taken at the point where the true depth is 2 feet will be negative?

3. Researchers at a large health maintenance organization (HMO) are planning a study of a certain mild illness. They will select a random sample of patients who are ages 35 to 54 and see if they contract the illness in the next year. The researchers are interested in estimating the proportion of men and of women who are likely to develop the illness in each of 4 age-groups: 35-39, 40-44, 45-49, and 50-54.
   The researcher plans to include 2000 patients in the study. Suppose the researchers draw a random sample from all of the patients at this HMO who are ages 35 to 54 and find the following numbers within each gender and age-group.

	Age-Group			
	35-39	40-44	45-49	50-54
Male	350	230	150	60
Female	445	370	245	150

a. Suppose that at the end of the study, 10 percent of the females in the 40-44 age-group contracted the illness. Calculate a 95 percent confidence interval to estimate the population proportion of females in this age-group that contracted the illness.

Interpret this confidence interval in the context of this situation.

Interpret the confidence level of 95 percent.

b. Suppose that at the end of the study, 10 percent of the males in the 40-44 age-group contracted the illness. The corresponding 95 percent confidence interval to estimate the population proportion of males in the age-group that contracted the illness is (0.061, 0.139).
Note that this interval and the interval in part (a) are of different lengths even though the two sample proportions were identical. What would be an alternative way to allocate a sample of 2000 subjects so that the 95 percent confidence interval widths for all male age-groups and for all female age-groups (i.e. for all 8 groups) would be the same when the sample proportions are the same? Justify your answer.

c. Based on previous studies, researchers believe that the percentages of those who contract the illness will be similar for males and females, and therefore plan to ignore gender when selecting a sample for this study. Previous studies also indicate the percentages of adults who will contract this illness in the 35-39, 40-44, 45-49, and 50-54 age-group are anticipated to be 5%, 8%, 20%, and 35%, respectively. How should the sample of 2000 subjects be allocated with respect to age-groups so that the widths of the 95 percent confidence intervals for the four groups will be approximately the same? Justify your answer.