

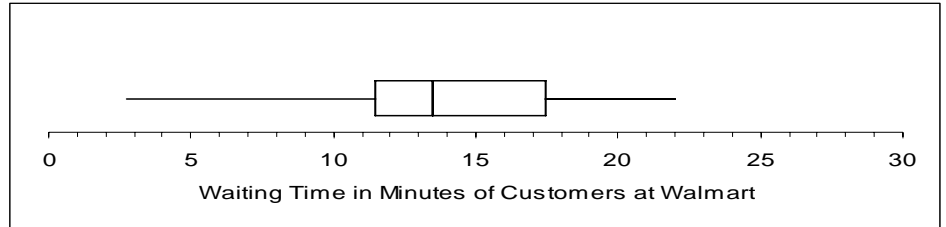


Pledge:

**Problem I.** Continued

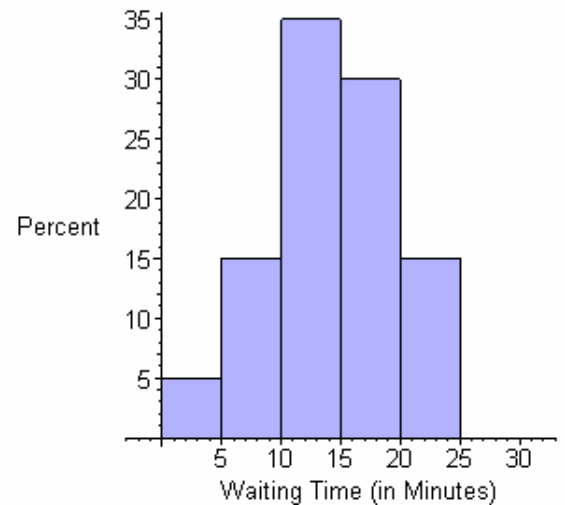
(f) The efficiency group also studied the waiting times of customers in check out lines at Walmart. Below you are given a boxplot and histogram of the Walmart data. Graph the boxplot of the waiting times for customers at Target above the boxplot of waiting times for Walmart. (5 points)

(g) Which group of customers (Target or Walmart) had the longest waiting time? What is that time? (2 points)



(h) Without doing any computations, which group of customers (Target or Walmart) had the largest *average* waiting time? (2 points)

Histogram of Waiting Times for Customers at Walmart



**Problem II.** You are contemplating buying a home. For each home you consider, you keep track of the variables below. Please indicate in the blank next to each variable whether it is quantitative or qualitative. (1 point each, 5 total)

- \_\_\_\_\_ Square feet
- \_\_\_\_\_ School district in which it is located
- \_\_\_\_\_ Style (Ranch, cottage, 2 story, etc.)
- \_\_\_\_\_ Number of Bedrooms
- \_\_\_\_\_ Price (in dollars)

**Problem III.** Please write “true” or “false,” according to which is correct, in the blank provided next to each statement. (2 points each, 10 points total)

- \_\_\_\_\_ The median is more sensitive to outlier data than the mean.
- \_\_\_\_\_ If the sign of the correlation coefficient is positive then the two variables from which it is computed have a positive association.
- \_\_\_\_\_  $0 \leq r^2 \leq 1$
- \_\_\_\_\_ Data from a scientifically conducted opinion poll is an example of experimental data.
- \_\_\_\_\_ If two variables have a strong positive association then we can say that the explanatory variable causes the response variable.

Pledge:

**Problem IV.** Suppose the distribution of ACT scores is approximately normal with mean 20.8 and standard deviation 4.8 and that the distribution of SAT scores are also approximately normal with mean 1026 and standard deviation 209. Use this information to answer the following questions: (10 points total)

(a) If Susie scores a 900 on the SAT and Joey scores an 18 on the ACT, which student had the better score? Please show all work to justify your answer. (5 points)

(b) If Brandon had a standardized score (i.e. a z-score) of 1.32 on the ACT, then what was his original score on the ACT? Please show all work. (5 points)

**Problem V.** Suppose the amount of time to wait in line before reaching the cashier at Walmart is approximately normally distributed with a mean of 15 minutes and a standard deviation of 5 minutes. Please answer the following: (15 points total)

(a) Draw a graph of the distribution of waiting times at Walmart according to this model. Please be sure to label your horizontal axis and to show values  $\pm 1$ ,  $\pm 2$ , and  $\pm 3$  standard deviations from the mean. (5 points)

(b) Using this model, determine what percent of customers at Walmart wait at most 8 minutes before reaching a cashier. Represent this percent on your graph above. (5 points)

(c) What are the smallest and largest times that determine the middle 80% of waiting times? (5 points)

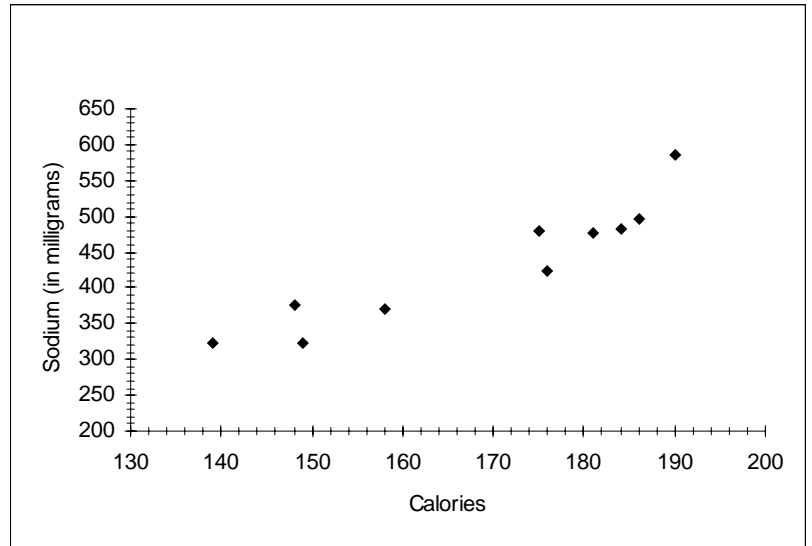
Pledge:

**Problem VI.** Consumer Reports Magazine measured the caloric content and sodium content (in milligrams) for 10 beef hot dogs. The two variables (caloric content and sodium) have a correlation coefficient of 0.9332. Below you are given a scatterplot of the data as well as some descriptive statistics of the data. Please answer the following questions. (24 points total)

Statistic	Calories	Sodium Content (in Milligrams)
Sample Mean	168.60	433.40
Sample Standard Deviation	18.39	85.67

(a) Plot the centroid of the data on the scatterplot. Clearly indicate the point and its coordinates on the plot. (2 points)

(b) By using the appropriate formulas, show that the equation of the regression line of Sodium on Calories is given by:  
$$\text{Sodium} = 4.347 * \text{Calories} - 299.504$$
  
Clearly show all of your work. (6 points)



(c) Accurately plot the regression line given in part (b) on the scatterplot below. Note that the vertical axis is at 130 calories. (4 points)

(d) Use the regression line to predict the amount of sodium, on average, for hotdogs that have a caloric content of 160 calories. Show this prediction graphically (i.e. using the up/down and over lines) on the scatterplot. Clearly indicate your answer below. (3 points)

(e) Consider the data point (190,587). Find the residual (i.e. prediction error) for this data point. (3 points)

(f) Using the regression model, what percent of the variation in sodium content is explained by caloric content? (2 points)

(g) Via this regression model, an increase in caloric content of 10 calories corresponds to what change in sodium content (i.e. an increase or decrease and by how much)? (2 points)

(h) Does caloric content cause sodium content in hotdogs? Can you think of some other explanation for the association between these two variables? (2 points)