

Pledge:

2/21/2006 MATH 171 – Statistical Decision Making Name: \_\_\_\_\_  
Dr. Lunsford Test 1 100 Points Possible

**Problem I.** The National Center for Education Statistics reported 1999 average mathematics achievement scores for eighth graders in 38 nations. Below is a frequency chart of the data. Please answer the following questions: (21 points total)

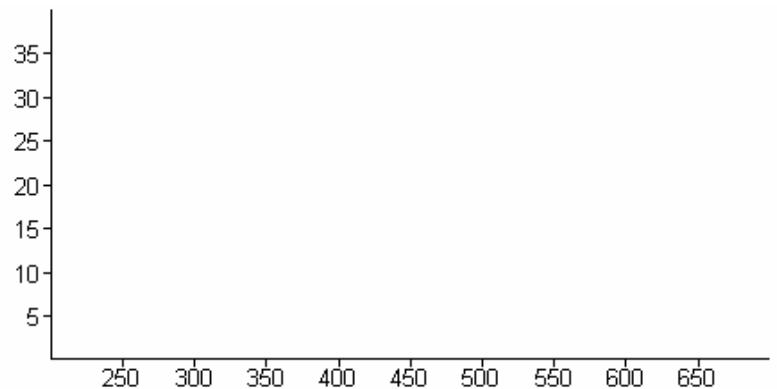
$X$ : Score	Frequency	Percent Frequency
$250 \leq X < 300$	1	
$300 \leq X < 350$	2	
$350 \leq X < 400$	1	
$400 \leq X < 450$	6	
$450 \leq X < 500$	9	
$500 \leq X < 550$	13	
$550 \leq X < 600$	5	
$600 \leq X < 650$	1	

(a) Please complete the frequency chart by finding the percent frequencies of the data. Only show the percent frequency values to one decimal place (i.e. xx.x%). (5 points)

(b) Use the frequency chart of the data to construct a percent frequency histogram on the axes provided below. Although I have provided values on the axes, please clearly label your axes for this histogram. (6 points)

(c) What percent of the 38 nations had an average achievement score of at least 550? Please show this percent graphically by shading the appropriate area on your histogram. (4 points)

(d) The United States had an average achievement score of 502. Which of the following best describes the performance of the United States eighth graders compared to their peers around the world (circle one)? (2 points)



Below average      About average      Above average

(e) Which is the best description of the shape of the distribution of average achievement scores for the 38 nations (circle one)? (2 points)

Left Skewed      Right Skewed      Uniform      Symmetric Unimodal      Bimodal

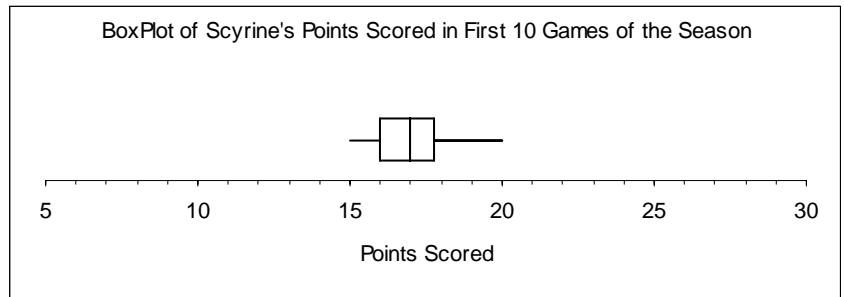
(f) For the distribution of average achievement scores for the 38 nations, which do you think is larger, the mean or the median? Why? (2 points)

Pledge:

**Problem II.** The basketball coach has kept statistics on her players throughout the semester. Below is a boxplot of the points scored during the first 10 games of the season by Scyrine. Also given below are the points scored during the first 10 games of the season by Alexandra. Please answer the following questions: (18 points total)

Points Scored by Alexandra (in ascending order): 6, 9, 11, 12, 17, 17, 20, 24, 26, 27

(a) Find the five number summary of Alexandra's points scored during the first 10 games of the season. Clearly indicate your answers. (5 points)



(b) Draw a box plot of Alexandra's points scored in the first 10 games of the season on the same axes as Scyrine's box plot (please draw Alexandra's above Scyrine's). (5 points)

(c) Do you think Alexandra and Scyrine score about the same number of points per game on average? Why or why not? You should answer this based purely on the box plots above, i.e. without doing any computations. (3 points)

(d) Which player has the smallest standard deviation in her points scored per game? Why? Again, you should answer this based purely on the box plots above, i.e. without doing any computations. (3 points)

(e) The coach can only take one of these players to the state championship. Which one should she take and why? (2 points)

**Problem III.** The State Education Department requires local school districts to keep certain records on all students. Below are some of the variables recorded for each student. Please indicate in the blank next to each variable whether it is quantitative or qualitative. (1 point each, 4 total)

- \_\_\_\_\_ age
- \_\_\_\_\_ race or ethnicity
- \_\_\_\_\_ standardized test scores in reading and mathematics
- \_\_\_\_\_ any disabilities or special needs the student may have

Pledge:

**Problem IV.** A clerk entering salary data into a company spreadsheet accidentally put an extra zero in the boss's salary, listing it as \$2,000,000 instead of \$200,000. For each of the following summary statistics, state how this mistake will affect the value of the statistic, i.e. will it cause a large change in the value or will it not cause much change? If it causes a large change, state how the incorrect value of the statistic will compare to the correct value (i.e. larger or smaller than the correct value). (2 points each, 8 total)

Mean:

Median:

Standard Deviation:

Quartiles:

**Problem V.** The Virginia Cooperative Extension reports that the mean weight of yearling Angus steers is 1152 pounds. Suppose that the weights of all such animals can be described by a normal model with a standard deviation of 84 pounds. Please answer the following questions. (19 points total)

(a) Draw a graph of the distribution of the weights of yearling Angus steers using the normal 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) What is the standardized score (i.e. z-score) for a steer weighing 1000 pounds? (3 points)

(c) What weight corresponds to a standardized score of 1.35? (3 points)

(d) Using this model, determine what percent of steers weigh at most 1000 pounds. Represent this percent on your graph above. (4 points)

(e) What are the smallest and largest weights that determine the middle 85% of weights? (4 points)

Pledge:

**Problem VI.** It is difficult to accurately determine a person's body fat percentage without immersing him or her in water. Researchers hoping to find ways to make a good estimate immersed 20 male subjects in water, then measured their waists and recorded their weights. Below are the descriptive statistics of the waist measurements and body fat data. A scatterplot of body fat versus waist measurement is also given. Please answer the following questions: (30 points total)

(a) Plot the centroid of the data on the scatterplot. (2 points)

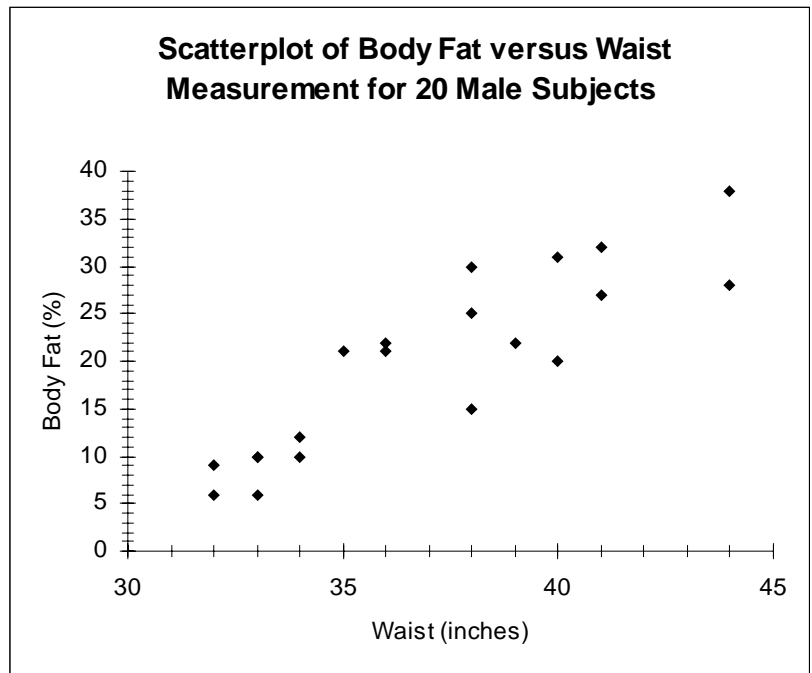
<b>Descriptive Statistics</b>	<b>Waist (in)</b>	<b>Body Fat (%)</b>
mean	37.05	19.75
standard deviation	3.82	9.56

(b) Researchers found a correlation coefficient of  $r = 0.8868$  for the variables body fat and waist measurement. Find the equation of the regression line of body fat on waist measurement. Neatly show all of your work. (8 points)

(c) Accurately plot the regression line you found in part (b) on the scatterplot below. Note that the vertical axis is at 30 inches. (4 points)

(d) Use the regression line to predict the percent body fat, on average, for men with a waist measurement of 36 inches. Show this prediction graphically on the scatterplot below. (3 points)

(e) Consider the data point (38,15). Find the residual (i.e. prediction error) for this data point. (3 points)



(f) What percent of the variation in body fat is explained by waist measurement? (4 points)

(g) For each increase of one inch in waist measurement we can expect what change in body fat (i.e. an increase or decrease and by how much)? (3 points)

(h) The researchers found the correlation between body fat and weight to be  $r = 0.6966$ . Based on this, which variable, weight or waist measurement is a better predictor (using a linear model) of percent body fat? Why? (3 points)