1. (*5 points*) The use of a regression line to predict outside the range of your data set that you used to obtain the line is called

A. interpretation

B. intrapolation

C. interpolation

D. extrapolation

E. mediation

2. (*5 points*) Using a regression to make a prediction for a value which falls far outside the range of your data set

A. is discouraged if we cannot conclude a cause and effect relationship

B. often gives inaccurate and/or absurd results

C. is considered good statistical practice only when the effect of lurking variables has been accounted for

D. very often gives accurate results

3. (*5 points*) Correlations based on averages can be misleading if

A. the relationship is not directly cause and effect.

B. they are very small.

C. they are interpreted to be about individuals.

D. they are very large.

E. there are potential lurking variables.

4. (*5 points*) For a biology project, you measure the weight in grams and the tail length in millimeters of a group of mice. The correlation is $r = 0.7$. If you had measured tail length in centimeters instead of millimeters, what would be the correlation? (There are 10 millimeters in a centimeter.)

A. 0.7

B. (0.7)(10) = 7

C. 0.7/10 = 0.07

5. (*20 points*) Calculate by hand the standard deviation of the data set: 0.67, 4.46, and 2.61. (To receive credit you must show detailed and well written calculations of both the mean and standard deviation.)

6. (*20 points*) An exercise in our textbook gives data on beak heat loss, as a percent of total body heat loss from all sources, at various temperatures. The data show that beak heat loss is higher at higher temperatures and that the relationship is roughly linear. The figure below shows Minitab regression output for these data.

Regression Analysis: Percent heat loss versus Temperature

The regression equation is
Percent heat loss = 0.92 + 2.06 Temperature

Predictor	Coef	SE Coef	T	P
Constant	0.919	5.613	0.16	0.872
distress	2.0647	0.2444	8.45	0.111

S = 4.50655 R-Sq = 83.6% R-Sq(adj) = 82.4%

(a) What is the equation of the least-squares regression line for predicting beak heat loss, as a percent of total body heat loss from all sources, from temperature? Use the equation to predict beak heat loss, as a percent of total body heat loss from all sources, at a temperature of 25 degrees Celsius.

(b) What percent of the variation in beak heat loss is explained by the straight-line relationship with temperature?

(c) Use the information in the above figure to find the correlation r between beak heat loss and temperature. How do you know whether the sign of r is $+$ or $-$?

7. (*20 points*) Emissions of sulfur dioxide by industry set off chemical changes in the atmosphere that result in “acid rain.” The acidity of liquids is measured by pH on a scale of 0 to 14. Distilled water has pH 7.0, and lower pH values indicate acidity. Normal rain is somewhat acidic, so acid rain is sometimes defined as rainfall with a pH below 5.0. The pH of rain at one location varies among rainy days according to a Normal distribution with mean 5.43 and standard deviation 0.54. What proportion of rainy days have rainfall with pH below 5.0?

8. (*20 points*) Researchers in New Zealand interviewed 907 drivers at age 21. They had data on traffic accidents and they asked the drivers about marijuana use. Here are data on the numbers of accidents caused by these drivers at age 19, broken down by marijuana use at the same age:

	Never	1-10 times	11-50 times	51+ times
Drivers	452	229	70	156
Accidents Caused	59	36	15	50

(a) Explain carefully why a useful graph must compare rates (accidents per driver) rather than counts of accidents in the four marijuana use classes.

(b) Compute the accident rates in the four marijuana use classes. After you have done this, make a graph that displays the accident rate for each class. What do you conclude? (You can't conclude that marijuana use causes accidents, because risk takers are more likely both to drive aggressively and to use marijuana.)