Classroom Activity 2 Math 113 Name : _____
10 pts Intro to Applied Stats

**Instructions:**

Is your pulse rate related to your blood pressure? Find your blood pressure and pulse rates and record them.

1. Record your blood pressure, and pulse rate in the table.

| Systolic Blood Pressure | Diastolic Blood Pressure | Pulse Rate |
|---|---|---|
| | | |

Give your information to the instructor who will enter it into Minitab and make a file available for the class to use.

2. Describe the who, what, where, when, why, and how of the data.

3. Blood pressure is usually measured in torrs (previously known as "mm Hg" or millimeters of mercury) and pulse rate in bpm (beats per minute). Some of you will be asked to convert units and see what effect that has on the correlation.

   Here are some common pressure conversion factors. Since they are given as 1 torr equals something, you multiply by the conversion factor to convert to that scale.

   1 torr = 0.0193368 pounds per square inch (psi)
   1 torr = 133.3224 pascals (Pa)

   Our group will measure the blood pressure in ( torrs / psi / pascals )

Convert any units before continuing with the activity.

4. Each group may be asked to work with a different set of predictor (x) and response (y) variables. Circle the appropriate variables for your group.

   Our predictor variable is ( systolic pressure / diastolic pressure / pulse rate ).
   Our response variable is ( systolic pressure / diastolic pressure / pulse rate ).

5. Generate a scatter plot of the response (y) and predictor (x) variables. Based on only the scatter plot, try to estimate the value of the correlation coefficient.

   Estimate of correlation coefficient, r = _____

6. Standardize the two variables and make a fitted line plot of the standardized variables. Find two points on the line and estimate the slope of the line as an estimate of the correlation coefficient, r.

   Slope of line as estimate of r = _____

7. Summarize the data. Circle the variable names and proper units so it is clear what you are describing. Indicate whether each variable is the predictor variable or the response variable.

| | **Predictor Variable** | **Response Variable** |
|---|---|---|
| **variable** | systolic / diastolic / pulse | systolic / diastolic / pulse |
| **units** | torrs / psi / pascals / bpm | torrs / psi / pascals / bpm |
| **sample size** | | |
| **mean** | | |
| **st. dev** | | |

8. Perform correlation between your two variables to find the actual value for the correlation coefficient, r. The correlation output will give you two numbers, a correlation coefficient and a probability value.

   Calculated correlation coefficient, r = _____

   Probability value, p-value = _____

9. Perform correlation again, but this time switch the order of the two variables. Record the correlation coefficient and p-value.

Calculated correlation coefficient, r = _____

Probability value, p-value = _____

What can you conclude about the order of the variables when finding correlation?

10. Talk to other groups and collect their correlation coefficients to complete as much of these tables as possible. Then, go back and fill in the any other numbers based on what you've learned of transforming the data.

**Table of correlation coefficients, r**

| | systolic (torr) | systolic (psi) | systolic (Pa) | pulse rate |
|---|---|---|---|---|
| diastolic (torr) | | | | |
| diastolic (psi) | | | | |
| diastolic (Pa) | | | | |
| pulse rate | | | | 1.000 |

What can you conclude about changing the scale on the variables?

11. The probability-value is the chance of getting the results we did if there is no linear correlation between the two variables. A small probability value (usually less than 5% or 0.05) means that these results are unlikely if there is no correlation and in turn means that they are linearly correlated. Use the p-value you found in question 8.

Based on our p-value of _____, the two variables _____ and _____ ( do / do not ) seem to be linearly correlated to each other.

12. Use Minitab to perform Regression on your two variables. Pay careful attention to which variable is the response variable and which is the predictor variable.

    Write the regression equation given by the computer.

13. Copy down the ANOVA table from the regression

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | | | | | |
| Residual | | | | | |
| Total | | | | | |

Here is an explanation of ANOVA table.

Source The source of the variation/variance. The Regression row is due to the regression equation and represents the part that can be explained by the model. The Residual row represents the part that can not be explained. The Residual source is sometimes called the Error.

DF Degrees of Freedom. The df(Regression) is one less than the number of variables. In this problem, there are 2 variables, the predictor and the response, so there is only one df for the Regression source. The df(Residual) is the sample size minus the number of variables. The df(Total) is one less than the sample size.

SS Sum of Squares of the deviations from the mean. Also known as variation.

MS Mean of the Squared deviation. This is found by dividing the variation (SS) by degrees of freedom (df) and is also known as the variance.

F Test statistic. We'll explain this later.

P Probability-value corresponding to the test statistic F.

14. Answer the following questions based on the SS column from the ANOVA table.

a. What is the explained variation?

b. What is the total variation?

c. Find the coefficient of determination, $r^2$, by dividing the explained variation by the total variation.

d. Find *r* by taking the square root of the coefficient of determination.

e. Compare the result of part d to the correlation coefficient found in 8. Ignoring that *r* might be negative, the results ( do / do not ) agree.

f. Take the total variation and divide it by the total degrees of freedom. This is the overall variance.

g. Find the standard deviation by taking the square root of the variance.

h. How does this standard deviation compare to the standard deviation of the response variable in question 7?

15. How does the p-value from the ANOVA table compare with the p-value from the correlation in question 8?