Try to answer the following six practice questions in no more than 10 minutes.
We will review the answers on Thursday.

1. A questionnaire is administered to all students in an Introductory Psychology class and the answers are used to obtain scores on two variables: X and Y. The correlation between X and Y is -0.5. A scatterplot of the two variables is shown in Figure 1. Write the regression coefficients for the best-fitting (i.e., regression) line for data shown in the Figure 1 and draw the regression line through the points in figure. Make sure to label/name each coefficient in your answer.

   Answer:

   ![Scatter plot for question 1](image)

   Figure 1: Scatter plot for question 1.

2. In the case of linear regression, which of the following statements about Pearson ($r$) and the Multiple Correlation Coefficient ($R$) is true?

   (a) It is not possible to use $r$ to calculate the value of $R^2$.
   (b) The value of $R^2$ is a better measure of goodness-of-fit than $r^2$.
   (c) Both $R^2$ and $r^2$ equal the squared value of the slope of the regression ($\beta^2$) line fit to standard scores.
   (d) You can use $R$ but not $r$ to calculate the slope of the regression line.

   Answer:
3. The scatter plot in Figure 2 illustrates the relation between variables $X$ and $Y$. Given those data, what can be said about the Pearson correlation $r$ and Spearman’s rank-order correlation ($r_s$)?

(a) $r$ will be greater than zero and $r_s$ will equal 1.
(b) $r_s$ will be greater than zero, but $r$ will be approximately zero because $X$ and $Y$ are not linearly related.
(c) Both $r$ and $r_s$ will be approximately zero because $X$ and $Y$ are not linearly related.
(d) Both $r$ and $r_s$ will be between zero and 1.

**Answer:**

![Figure 2: Scatter plot for question 3.](image)

4. Simpson’s paradox

(a) is the title of a 1999 episode of The Simpson’s, in which Stephen Hawking discusses the idea of a donut-shaped universe with Homer, before punching him in the face with his robotic boxing glove.
(b) refers to the fact that the magnitude and direction of a correlation between two variables can differ markedly between sub-groups in a heterogeneous sample.
(c) is another name for “regression toward the mean”.
(d) refers to the fact that the IQR for a skewed distribution is always larger than the IQR for a symmetric distribution when the two distributions have the same variance.

**Answer:**

5. When an investigator says that a correlation is “significant ($p < .05$)”, she means

(a) the correlation has important theoretical and/or practical consequences.
(b) there is a high probability ($p \geq 0.95$) that the observed correlation was caused by chance.
(c) a correlation at least as large as the observed $r$ would be expected to occur by chance less than 5% of the time.
(d) the correlation is large.

**Answer:**
6. Even a high correlation between X and Y does not imply that X causes Y because
   (a) The high correlation could be obtained if Y caused X.
   (b) The correlation may reflect the fact that both X and Y are correlated with a third variable.
   (c) The correlation might be due to the fact that X and Y are simply changing over time.
   (d) All of the above are reasonable arguments against concluding that X caused Y.

   **Answer:** ________________________________