Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

AP STATISTICS – MS. KLIMCZUK

**Chapter 8 Questions (From Notes)**

1. Here is the scatterplot of *Dive Heart Rate* (in beats per minute) vs. *Duration* (in minutes) of emperor penguin.



1. Describe this graph.

**Details:**

**R² = 71.5%** $\hat{Dive Heart Rate}=96.9 -5.47 (Duration)$

1. Explain what the R² tells us.
2. What does the slope mean in the context of this problem?
3. What does the y-intercept mean?
4. Here is the scatterplot of residuals against *Duration* of emperor penguin. What does the residuals plot tell you about the appropriateness of using a linear model?



1. How are residual plots important?
2. Here is a histogram of the residuals in the regression predicting *Calories* from *Sugar* content in cereals.



1. How would you describe its shape? What do you notice from looking at the histogram?
2. What might be going on here?

Here is the residuals plot.



1. The small modes in the histogram are marked with different symbols in the residual plot. What do you see?
2. The high-residuals cereals are Just Right, Fruit & Nut, Muesli Raisins, Dates and Almonds, Peaches and Pecans, Nutri-Grain Almond Raisins, and Mueslix Crispy Blend. What do these cereals all have in common?
3. The low-residuals cereals are Puffed Rice, Puffed Wheat, three bran cereals, and Golden Crisps. What do these cereals all have in common?
4. What is extrapolation?
5. Here is a timeplot of the Energy Information Administration (EIA) predictions and actual prices of oil barrel prices.



1. How did forecasters do?
2. Do you think you can predict where oil prices will go in the next decade? Why or why not?
3. What is a high leverage point? Draw an example.
4. A point with high leverage has the potential to change the regression line, but it doesn’t always use that potential. What would happen if our high leverage point lined up with the other points? Draw an example.
5. What is an influential point? Draw an example.

**Just Checking Questions on Pg 217 (Write down answers below):**

1. What is a Lurking Variable?
2. The following scatterplot shows that the average *life expectancy* for a country is related to the number of *doctors* per person in that country:



The square root is used here to make sure the scatterplot fits the Straight Enough Condition.

The R² = 62.4%.

1. What does this graph make you believe?
2. The strength of the association would seem to argue that we should send more doctor’s to developing countries to increase life expectancy. Would this work?
3. Do doctor’s CAUSE greater life expectancy?
4. What could be a lurking variable here?
5. This new scatterplot shows that the average *life expectancy* for a country is related to the number of *televisions* per person in that country:



The association in this scatterplot is even stronger.

R² = 72.3%

1. Since televisions are cheaper than doctors, do you think sending TVs to countries with low life expectancies would extend lifetimes?
2. What is going on here?
3. Why would scatterplots of statistics summarized over groups tend to show less variability than we would see if we measured the same variable on individuals?
4. Describe this graph:



 R² = 41.5%

1. Describe this graph:



R² = 80.1%

1. Now compare the two graphs in Example 12 and 13.