Lesson #69 · Topic 5B · Practice

Name ____________________

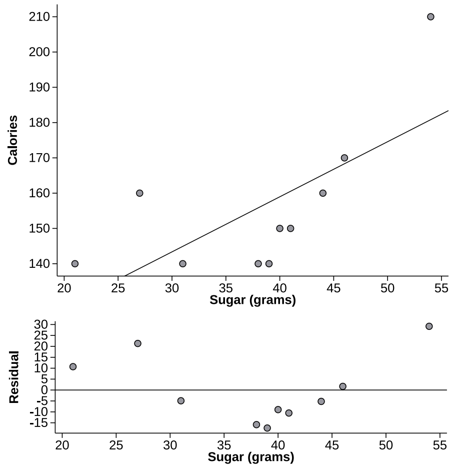
1. A country's natural resources, including oil, gas, and coal, can greatly impact the country's overall economy, as measured by gross domestic product. Listed in the table are the eight countries with the most revenue from natural resources, in billions of dollars, along with each country's gross domestic product, GDP, in billions of dollars.

Country	Natural resources ($ billions)	GDP ($ billions)
Kuwait	30.0	106
Azerbaijan	16.3	54.6
Angola	20.2	67.4
Iran	109.3	359.7
Mongolia	5.0	15.3
Democratic Republic of Congo	21.5	55.4
Iraq	90.2	207.9
Libya	26.1	42.8

a. Using technology, find the least-squares regression line relating the revenue from natural resources and the GDP. Be sure to define any variables used.

b. Use the least-squares regression line to predict the GDP for a country with a revenue of $50 billion from natural resources.

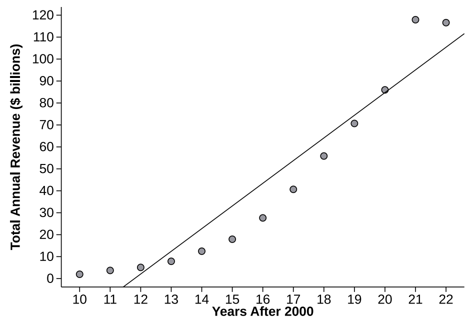
2. The amount of sugar, in grams, and the number of calories in a 12-ounce serving of 10 different popular soft drinks are shown in the scatterplot. The scatterplot with the least-squares regression line and the residual plot are shown.



a. The sum of the squared residuals is 2,219.9. Which point contributes the most to this sum? Explain.

b. Is a linear model appropriate for these data? Explain.

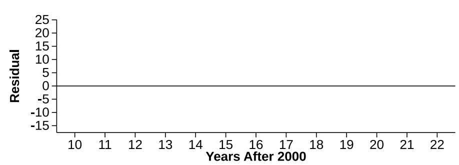
3. The scatterplot shows the total annual revenue (in billions of dollars) for the social media website Facebook, for each year from 2010 to 2022 recorded as years since 2000.



a. The least-squares regression line for the relationship is $\hat{y} = -121.89 + 10.33x$. Calculate and interpret the residual for the year 2016 if the actual revenue was $27.64 billion.

b. Without any calculations, use the scatterplot and the graph of the least squares regression line to estimate the residual for the year 2021.

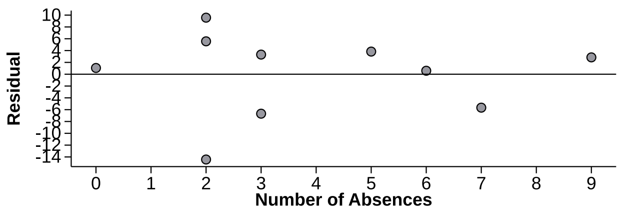
c. Sketch a residual plot for the data on the axes given by approximating each residual.



d. The correlation of the relationship is 0.95. Is a linear model a good fit for the data? Explain.

4. At the end of a semester, a math teacher wonders if student attendance has an impact on final exam scores. The teacher selects a random sample of 10 students and records the number of days absent and final exam score for each student. The equation of the least-squares regression line and a residual plot are provided.

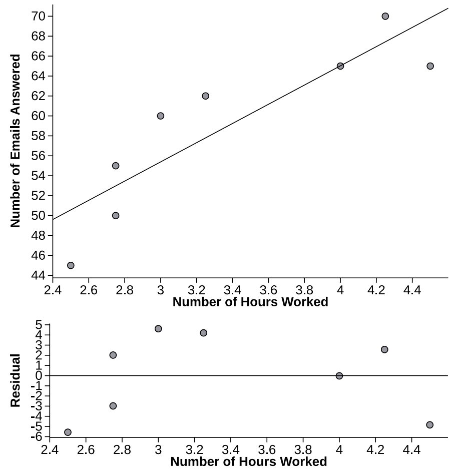
$$\widehat{\text{final exam score}} = 89.943 - 2.755(\text{number of days absent})$$



a. Is the linear model appropriate? Explain.

b. What is the actual exam score for the student with 5 absences? Round to the nearest whole point.

5. Emily works as a customer service representative answering customer emails. For eight days she records the amount of time she spends answering emails and the number of emails she can answer during that time. Her results are shown in the scatterplot with the least squares regression line and the residual plot.



a. The residual for 3 hours is 4.5. Interpret this value.

b. The residual for 4 hours is 0. Explain what this means.

c. Here is a list of each residual, rounded to the nearest tenth.

$$-5.6, -4.8, -3, 0, 2, 2.6, 4.2, 4.6$$

The mean of the residuals is 0 and the standard deviation of the residuals is 4.017. Interpret this standard deviation.

6. A student records the amount of protein (in grams) and calories per serving in 14 types of cheeses and butters. The equation for the least-squares regression line used to predict the amount of protein from the number of calories is:

$$\widehat{\text{grams of protein}} = 1.575 + 0.032(\text{number of calories})$$

One slice of American Pasteurized cheese has 79 calories and 4.7 grams of protein. Based on the residual, does the regression model overestimate or underestimate the amount of protein for this type of cheese?

A) Underestimate, because the residual is positive.

B) Underestimate, because the residual is negative.

C) Overestimate, because the residual is positive.

D) Overestimate, because the residual is negative.

7. Scientists compared the maximum weight (in pounds) and average life expectancy (in years) of 16 dog breeds. Let x be the average life expectancy of the dog breed and y be the maximum weight of the dog breed. The least-squares regression line $\hat{y} = 329.4 - 21.6x$ was calculated from the data. Suppose all the values for maximum dog weight were converted from pounds to kilograms (1 kilogram = 2.2 pounds). Which of the following gives the equation of the least-squares regression line for the updated data with $x =$ average life expectancy and $y =$ maximum weight (in kilograms)?

A) $\hat{y} = 149.73 - 21.6x$

B) $\hat{y} = 149.73 - 9.82x$

C) $\hat{y} = 329.4 - 9.82x$

D) $\hat{y} = 329.4 - 21.6x$