

Correlation and Regression

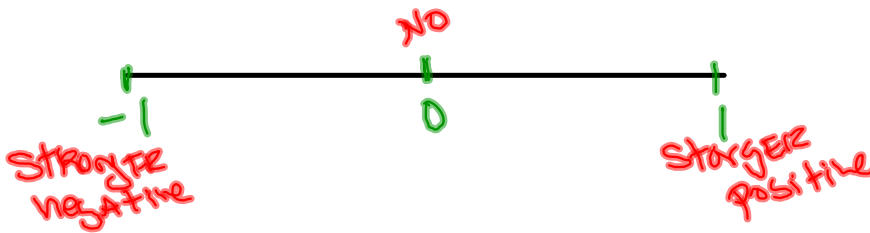
7. A study of 10 countries showed that smaller the percentage change in wages, smaller is the percentage change in consumer prices.

- a. Identify the explanatory variable and state its units
- b. Identify the response variable and state its units
- c. Which of the following is a possible value for the linear correlation coefficient between the percentage change in wages and the percentage change in consumer prices?

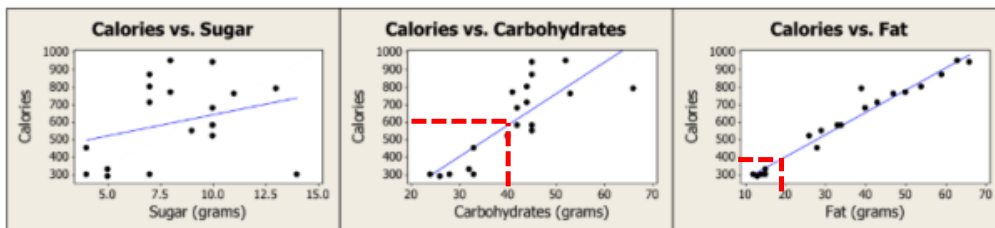
→ changes in wages (%)
→ changes in consumer prices (r?)

- A) 0.7 B) -0.65 C) -0.85 D) 0.1

(R) (r) correlation coefficient



8. Researchers wanted to study the relationship between amounts of fat, sugar, and carbohydrates and the amount of calories in a hamburger. They gathered relevant data about 22 “brands” of fast food hamburgers and obtained the following scatter plots:



- About how many calories would you predict for a burger that has 20 grams of fat?
- About how many calories would you predict for a hamburger that has 40 grams of carbohydrates?
- Which prediction is likely to be more accurate? Why do you think this?
- Which nutrient has the weakest impact on calories? Why do you think this?
- What does the idea of strength of the correlation tell you about whether a nutrient is a good predictor of calories?
- What is the direction of the fat/calories graph? What does the direction of the line tell you about the association between the amount of fat and the calories in fast food hamburgers?

400 cal.

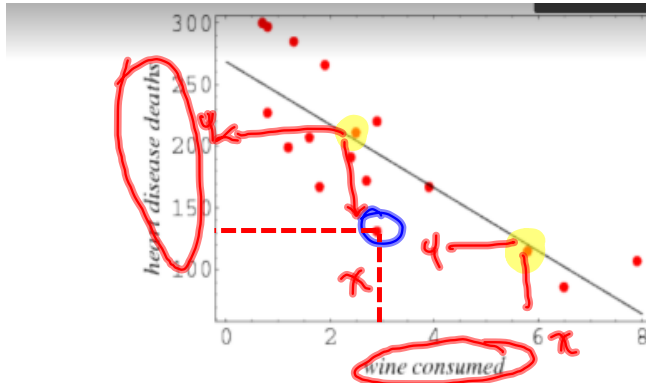
600 cal.

— CAL vs. Fat (they tend to form a line.)

→ SUGAR

up / positive

9. The scatterplot below relates **wine consumption** (in liters of alcohol from wine per person per year) and **death rate from heart disease** (in deaths per 100,000 people) for 19 developed countries.



$$(x_1, y_1) \quad (x_2, y_2)$$

$$y_2 - y_1 = m(x_2 - x_1)$$

- a. Identify an individual (subject) in this study? →
- b. What does each dot in this scatterplot represent? →
- c. What type of correlation is shown in this scatterplot? Circle the correct answer:
 Linear Non-linear No correlation
- d. What is the direction of association between variables? Circle the correct answer:
 Positive Negative None

For questions e. and f. use the equation of the Least-Square Regression LSR line is:

$$\hat{y} = -22.97x + 260.56 \quad \text{-model}$$

- e. Circle the correct choice and fill in the blank in the following statement: As wine consumption increases by 1 liter of alcohol per person per year, the predicted death Rate from heart disease increases decreases by _____ deaths per _____ people.
- f. Find the death rate from heart disease (per 100,000 people) predicted by the model for a country where wine consumption amounts to 5 liters of alcohol from wine per person per year?

increase by 1 liter ↓

$$y = -22.97(1) + 260.56 = 237.59$$

$$y = -22.97(2) + 260.56 = 214.62$$

$$\quad \quad \quad \checkmark \frac{214.62}{22.97}$$

$$y = -22.97(5) + 260.56$$

$$y = 145.71$$