

Regression Example 1

Alison Rector

2022-09-12

Running Example

Michael is a runner. He uses his Apple watch to keep track of his distance and the number of calories he burns for 20 runs. A scatterplot of y = calories burned and x = distance (in miles) shows a fairly strong, positive linear relationship. The regression equation $y = 20 + 160x$ models the data fairly well.

a. Interpret the slope of the regression line.

Michael burns 160 calories for every 1 mile that he runs.

b. Does the value of the y intercept have meaning in this context? If so, interpret the y intercept. If not, explain why.

Regardless of the number of miles that Michael runs, there is a baseline additional burn of 20 calories. However, this is not interpretable in the context of Michael running zero miles.

c. Predict the number of calories Michael burns if he runs 5 miles.

$$\hat{y} = 20 + 160(5) = 20 + 800 = 820$$

d. Calculate and interpret the residual if his Apple watch said that he burned 910 calories on a 5-mile run.

$$r_i = (y_i - \hat{y}) = (820 \text{ kcal} - 910 \text{ kcal}) = -90 \text{ kcal}$$

The predicted value for calories burned for a 5-mile run was 90 calories less than the observed value.

e. Michael is thinking about signing up for his first marathon. So far, his longest run has only been 10 miles. Should he use the regression equation to predict how many calories he would burn if he runs a marathon (26.2 miles)? Explain.

No, Michael should not use the regression equation because 26.2 miles is too far outside of the observed prediction values used to construct his model.