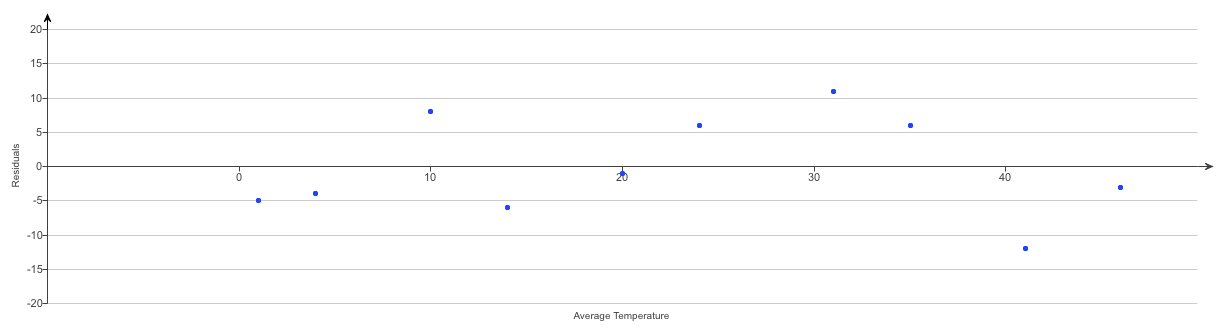
After keeping track of his heating expenses for several​ winters, a homeowner believes he can estimate the monthly cost from the average daily Fahrenheit temperature using the model Cost=129−1.95 *temp*. The residuals plot for his data is shown.

a) Interpret the slope of the line in this context. What does the model predict given a 1° increase in​ temperature?

**A.**

The model predicts a ​$129 increase in heating cost.

**B.**

The model predicts a ​$1.95 increase in heating cost.

**C.**

The model predicts a ​$129 decrease in heating cost.

**D.**

The model predicts a ​$1.95 decrease in heating cost.

​b) Interpret the​ y-intercept of the line in this context. What does the model predict given a temperature of 0°​?

**A.**

The model predicts a heating cost of​ $0.

**B.**

The model predicts a heating cost of ​$1.95.

**C.**

The model predicts a heating cost of ​$129.

​c) During the months when the temperature stays around freezing ​(32 degrees ° ​F), would you expect the cost predictions based on the model to be​ accurate, too​ low, or too​ high? Explain.

**A.**

Too low, because the residual is positive.

**B.**

Accurate, because the predicted and observed values are very close.

**C.**

Too high, because the residual is negative.

d) What heating cost does the model predict for a month that averages 10°​?

*cost=​$\_\_\_\_\_\_\_*

​(Round to the nearest cent as​ needed.)

​e) During one of the months on which the model is​ based, the temperature did average 10°. What were the actual heating costs for this​ month?

The actual heating costs for this month were ​$\_\_\_\_\_\_\_\_\_

​(Round to the nearest dollar as​ needed.)

f) Do you think the homeowner should use this​ model?

**A.**

No

**B.**

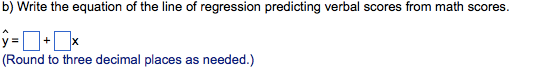
Yes

Question 3)

A group of high school seniors took a scholastic aptitude test. The resulting math scores had a mean 480.4 with a standard deviation of 190.1​, verbal scores had a mean 480.6 with a standard deviation of 162.3​, and the correlation between verbal and math scores was r=0.718. Complete parts a through f below.

a)What is the​ correlation?

The correlation is:\_\_\_\_\_\_\_\_\_



c) In​ general, what would a negative residual mean in this​ context?

1. a **negative residual means the student has**

a higher verbal score than than the linear model would predict.

**B.**

A negative residual means the student has

the exact verbal score that the linear model would predict.

**C.**

A negative residual means the student has a lower verbal score than the linear model would predict.

​d) A person tells you her math score was 241. Predict her verbal score.

The student is expected to have a verbal score of \_\_\_.

​(Round to three decimal places as​ needed.)

​e) Using the predicted verbal score from part​ (d) and the regression equation Math=76.215 +0.841(verbal)​, predict the​ student's math score.

The predicted math score is:\_\_\_\_\_\_.

​(Round to three decimal places as​ needed.)

​f) Why​ doesn't the result in part​ (e) come out to 241​?

**A.**

The line given in part​ (e) minimizes squared residuals in the verbal​ direction, while the line found in part​ (b) minimizes them in the math direction.

B.

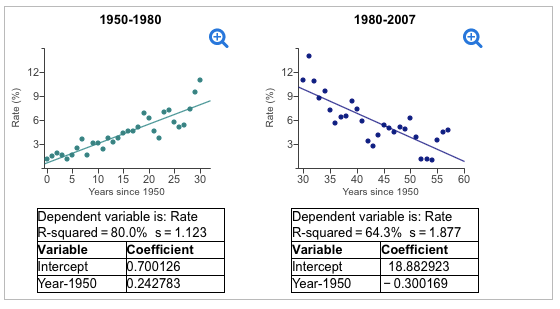
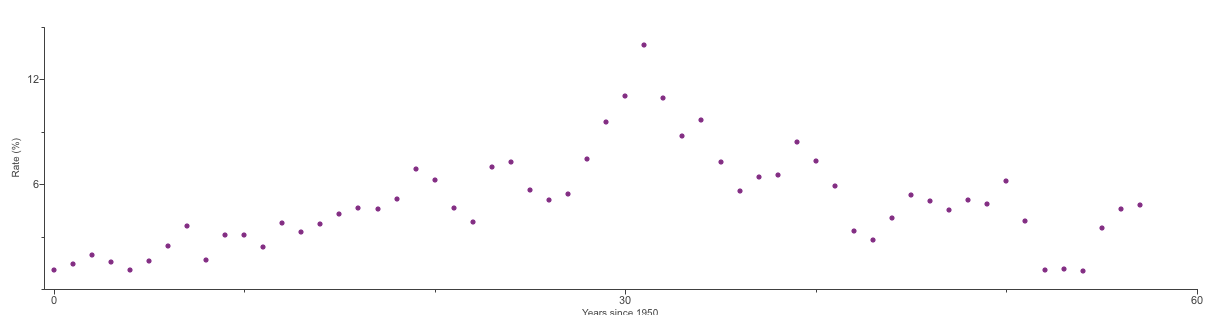
The line found in part​ (b) minimizes squared residuals in the math​ direction, while the line given in part​ (e) minimizes them in the verbal direction.

**C.**

The line found in part​ (b) minimizes squared residuals in the verbal​ direction, while the line given in part​ (e) minimizes them in the math direction.

question 4)

The scatterplot to the right shows that the trend for the interest rate on a​ 3-month bond changed dramatically after​ 1980, so two regression models were fit to the relationship between the rate​ (in %) and the number of years since​ 1950, one for 1950 to 1980 and one for the data from 1980 to 2007. The accompanying display shows the plots of the interest rate on the​ 3-month bond from 1950 to 1980 and from 1980 to 2007 and their corresponding regression models. Complete parts a through d.



​a) How does the model for the data between 1980 and 2007 compare to the one for the data between 1950 and​ 1980?

**A.**

The model for the data between 1980 and 2007 fits the data much better than the other model.

**B.**

The models are approximately equal.

**C.**

The two models both fit​ well, but they have very different slopes.

​b) What does the model for the data between 1980 and 2007 estimate the interest rate to have been in 1995​? How does this compare to the predicted rate obtained from the model for the data between 1950 and​ 1980, which is approximately 12​%?

The model estimates that the interest rate was:\_\_\_\_\_\_ ​% in 1995.

​(Round to one decimal place as​ needed.)

c) How does the predicted value from the model for the data between 1980 to 2007 compare to the predicted value from the model for the data between 1950 and​ 1980, which is approximately 12​%?

**A.**

The model for the data between 1980 and 2007 predicts the interest rate in 1995 to be much higher than the other model predicts.

**B.**

The model for the data between 1980 and 2007 predicts the interest rate in 1995 to be about the same as the other model predicts.

**C.**

The model for the data between 1980 and 2007 predicts the interest rate in 1995 to be much lower than the other model predicts.

​c) Do you trust this newer predicted​ value?

A.

Yes. R2=64.3% is large enough for the extrapolation to be safe

**B.**

Not really. Extrapolating 45 years beyond the beginning of these data would be dangerous and unlikely to be accurate.

**C.**

​Yes, because the​ x-value is within the range of the original data.

**D.**

Not really. Interest rates are too random to predict for any time.

​d) Given these two​ models, what would you predict the interest rate on the​ 3-month bond will be in 2020​?

**A.**

It would be best to take the mean of the predicted values from each​ model, which is about 8​%.

**B.**

It would be best to use the predicted value from the model for the data between 1980 and​ 2007, which is about 2​%.

**C.**

It would be best to use the predicted value from the model for the data between 1950 and​ 1980, which is about 18​%.

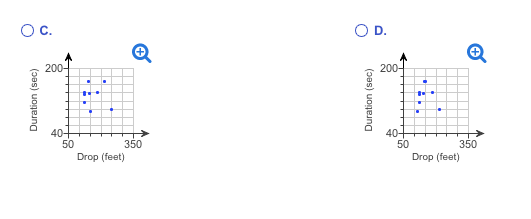
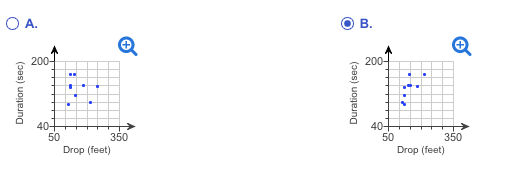
**D.**

It would be best not to predict the value because extrapolating beyond the​ x-values that were used to fit the model can be dangerous.

Question 5)

The accompanying table shows the height of the initial drop​ (in feet) and the duration of the ride​ (in seconds) of the 10 best roller coasters in a certain country. What do these data indicate about the height of the initial drop of a roller coaster and the duration of the​ ride, using correlation and a​ scatterplot?

Construct a scatterplot of the data. Choose the correct answer below. which is the correct scatter plot?



B) Calculate the correlation coefficient.

r=\_\_\_\_\_\_

​(Round to three decimal places as​ needed.)

c) What do these data indicate about the height of the initial drop of a roller coaster and the duration of the​ ride?

**A.**

​Generally, rides on coasters with a greater initial drop tend to last

somewhat shorter​, because the correlation indicates a weak association.

**B.**

​Generally, rides on coasters with a greater initial drop tend to last

somewhat longer​, because the correlation indicates a moderate association.

C.

​Generally, rides on coasters with a greater initial drop tend to last

longer, because the correlation indicates a strong association.

**D.**

​Generally, rides on coasters with a greater initial drop tend to last somewhat longer​, because the correlation indicates a weak association.