

Chapters 7-9 Review

Name: LEY

1. At Ginormous University, a statistics student gathered information on the amount of sleep the average student gets based on the number of hours they studied that evening. There were 116 students who responded to the questionnaire. The following table is the information computed by statistical software from the data.

Predictor	Coefficient	SE Coefficient
Constant	7.5555	0.2239
Hours of Studying	-0.2692	0.0662
$S_x=1.509$	$R^2=12.7\%$	

- a. Using the output from the table above, write the equation for the Least Squares Regression Line (Line of Best Fit)

$$\hat{y} = 7.5555 - 0.2692x$$

$x = \text{HOURS OF STUDYING}$   
 $\hat{y} = \text{HOURS OF SLEEP}$

- b. What is the value of the correlation coefficient? How do you know the sign ( $\pm$ ) of  $r$ ?

$$R = \sqrt{0.127}$$

SLOPE OF  $\hat{y}$  IS NEGATIVE

$$= -0.356$$

- c. What is the value of the coefficient of determination? Interpret this in the context of the problem.

$$R^2 = 0.127$$

$$12.7\%$$

- d. What is the predicted amount of sleep for a college student who studied 2.5 hours yesterday?

$$\hat{y} = 7.5555 - 0.2692(2.5)$$

$$= 6.883 \text{ HOURS}$$

- e. If a college student studied 5 hours and got 4.5 hours of sleep, what is the residual? Did the Least Square Regression Line over estimate or under estimate the amount of sleep?

$$(5, 4.5)$$

$$\hat{y} = 7.5555 - 0.2692(5)$$

$$\text{RESIDUAL} = 4.5 - 6.21$$

$$= 6.21$$

$$= -1.71$$

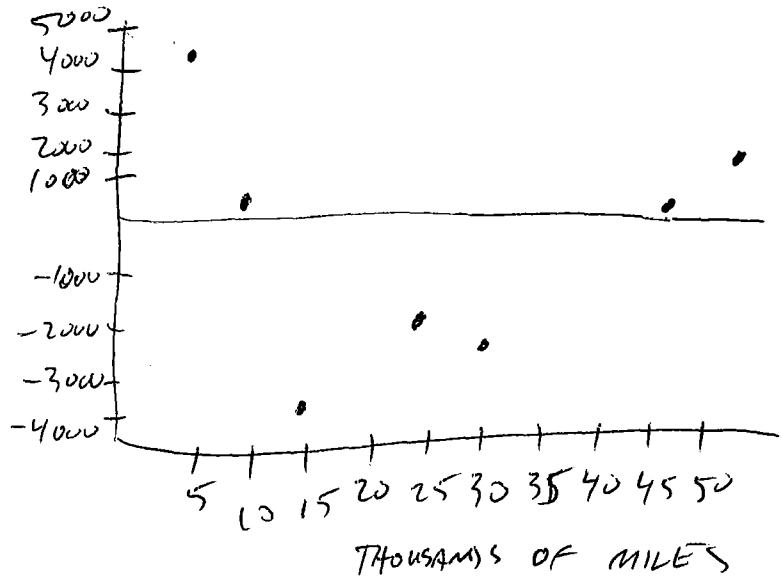
2. As more miles are driven in a car, the resale value of a car generally declines. This is called depreciation. The following data is for a certain make and model of car for the resale price in dollars and the number of miles driven (in thousands of miles).

Miles	5	9	15	24	30	45	50
Resale Price	32,500	27,000	21,000	19,500	17,000	14,000	13,900

Find the equation for the Least Squares Regression Line (Line of Best Fit) and sketch the Residual plot:

$$\hat{y} = 30010 - 366.1x$$

$\hat{y}$  = PREDICTED VALUE (PRICE)  
 $x$  = MILES



Estimate the resale price for a car with 23,000 miles:

$$30010 - 366.1(23) = 21,589.70$$

Multiple Choice. Chose the BEST answer:

3. A regression line was used to predict the final exam score based the first test score of the semester. The equation of the Least Squares Regression Line is:  $\hat{y} = 10 + 0.9x$  What represents the first test score?
- The Intercept
  - The Slope
  - The Explanatory Variable
  - The Response Variable
  - The Correlation Coefficient

4. The value of a correlation coefficient is a resistant statistic (not sensitive to outliers).

True

False

5. A Least Squares Regression Line was fitted to the weights (in kilograms) of a group of children between the ages of 12 and 36 months. The equation of the line is:  $Weight = 6.9 + 0.27Age$ . Weight is in kilograms and Age is in months. Which of the following is true?

- a. The predicted weight for a two year old is 7.44 kilograms
- b. The predicted weight for a two year old is 13.38 kilograms
- c. It is not appropriate to use this equation for a two year old since 2 is not between 12 and 36
- d. The predicted weight of a baby birth is about 6.9 kilograms
- e. None of the above are true

6. A Least Squares Regression Line was calculated for collected data is:  $\hat{y} = 17 + 2x$ . What is the predicted value when  $x=3$ ?

- a. 2
- b. 17
- c. 21
- d. 22
- e. 23

7. A Least Squares Regression Line was calculated for collected data is:  $\hat{y} = 17 + 2x$ . What is the value of the residual for the point (3, 22)?

- a. -2
- b. -1
- c. 0
- d. 1
- e. 2

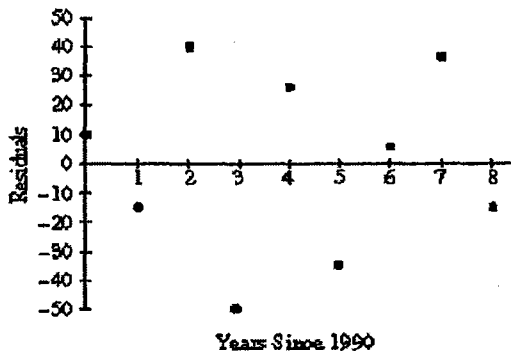
8. The Least Squares Regression Line is fit to a set of data. If one of the data points has a positive residual value, then what is known?

- a. The correlation between the values of the response and the explanatory variables must be positive.
- b. The point must lie above the Least Squares Regression Line.
- c. The point must lie below the Least Squares Regression Line.
- d. The point must lie near the right edge of the scatterplot.
- e. The point must lie on the Least Squares Regression Line.

9. What is the purpose of a Residual Plot?

- a. To determine causation
- b. To assess the type of relationship that exists between the explanatory and response variables
- c. To provide predictions for the explanatory variable
- d. To provide predictions for the response variable
- e. To check the appropriateness and fit of the Least Squares Regression Line for the given data

10. Some AP Statistics students were searching the Internet to find information on air travel in the United States. They found data on the number of commercial aircraft flying in the United States during the years 1990–1998. The dates were recorded as years since 1990. Thus, the year 1990 was recorded as year 0. They fit a regression line to the data. The graph of the residuals and part of the computer output for their regression are given below.



Predictor	Coef	StDev	t-ratio	p
Constant	2939.93	20.55	143.09	0.000
Years	233.517	4.316	54.11	0.000

s = 33.43

- (a) Is a line an appropriate model to use for these data (Is Linear Regression appropriate)? What information tells you this?

YES. THE RESIDUAL PLOT SHOWS NO PATTERN, THEREFORE A LINEAR MODEL IS APPROPRIATE

- (b) What is the value of the slope of the least squares regression line? Interpret the slope in the context of this situation.

$$\text{SLOPE} = 233,517$$

AVERAGE NUMBER OF COMMERCIAL AIRCRAFT FLYING IN U.S. INCREASING PER YEAR.

- (c) What is the value of the intercept of the least squares regression line? Interpret the intercept in the context of this situation.

$$\text{INTERCEPT} = 2939.93$$

PREDICTED NUMBER OF COMMERCIAL AIRCRAFT FLYING IN 1990  
(X = 0 CORRESPONDS TO 1990)

- (d) What is the predicted number of commercial aircraft flying in 1992?

$$1992 = \text{YEAR } 2.$$

$$3407$$

$$2939.93 + 233.517(2) = 3406.964 \text{ AIRCRAFT}$$

- (e) What was the actual number of commercial aircraft flying in 1992?

FROM THE RESIDUAL PLOT: RESIDUAL FOR 1992 IS +40

$$40 = \text{ACTUAL} - \text{PREDICTED}$$

$$40 = \text{ACTUAL} - 3407$$

$$\text{ACTUAL} = 3447$$