Bell Ringer

Section 6.4 Calculator

1. Use a graphing calculator to find the equation of the line of best fit for the data in the table. Find the value of the correlation coefficient r to three decimal places. Then predict the number of movie tickets sold in the U.S. in 2014.

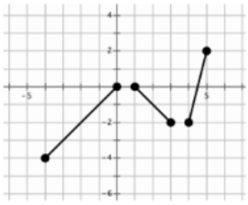
Movie Tickets Sold in U.S. by Year

x L_1 Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
y L_2 Tickets Sold (millions)	1289	1311	1340	1339	1406	1421	1470	1415	1472	1470

Source: Motion Picture Association of America

Review.

2. What is the Domain and Range of the function?



Solutions

Section 6.4 Calculator

1. Use a graphing calculator to find the equation of the line of best fit for the data in the table. Find the value of the correlation coefficient r to three decimal places. Then predict the number of movie tickets sold in the U.S. in 2014.

Movie Tickets Sold in U.S. by Year

x Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
y Tickets Sold (millions)	1289	1311	1340	1339	1406	1421	1470	1415	1472	1470

Source: Motion Picture Association of America

2014

y = 21.4x – 41,557

r = .942

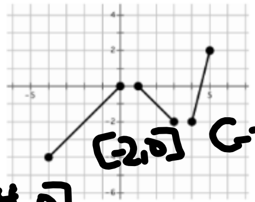
In 2014 about 1542.6 million tickets sold

1,542.6 million

21.4(2014) - 41,557

Review.

2. What is the Domain and Range of the function?



[-4,0] [-2,0] [-2,2]

Domain: [-4, 0] or [1, 3] or [4, 5]

Range: [-4, 2]

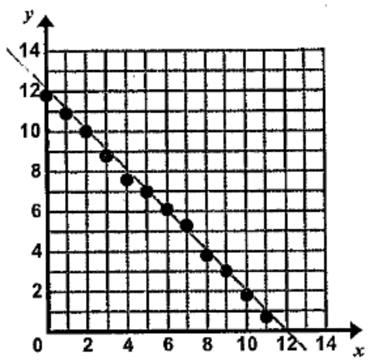
Correct 6.4A Trendlines ws

9.2 Assignment: Lines of Fit

Name Key Date__________ Period__________

Directions: In #1 and 2, observe the data sets and take note of any associations you see, draw a line of best fit, write a prediction function, and use your function to predict the value of y when $x = 12$ and when $x = 100$.

1.



a. Observations:

negative, linear

b. Using a ruler, draw a line of best fit through the data points that captures the general trend of the data.

c. Estimate the slope and y-intercept of your line.

$m \approx -1$ $b \approx 12$

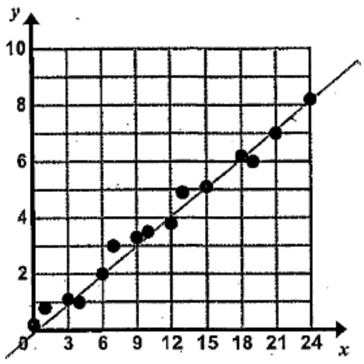
d. Write a prediction function for the data set. $y = -1x + 12$

e. Use your prediction function to find the value of y when $x = 12$ and when $x = 100$.

12: $y = 0$

100: $y = -88$

2.



a. Observations:

pos, linear

b. Using a ruler, draw a line of best fit through the data points that captures the general trend of the data.

c. Estimate the slope and y-intercept of your line.

$m \approx \frac{1}{3}$ $b \approx 0$

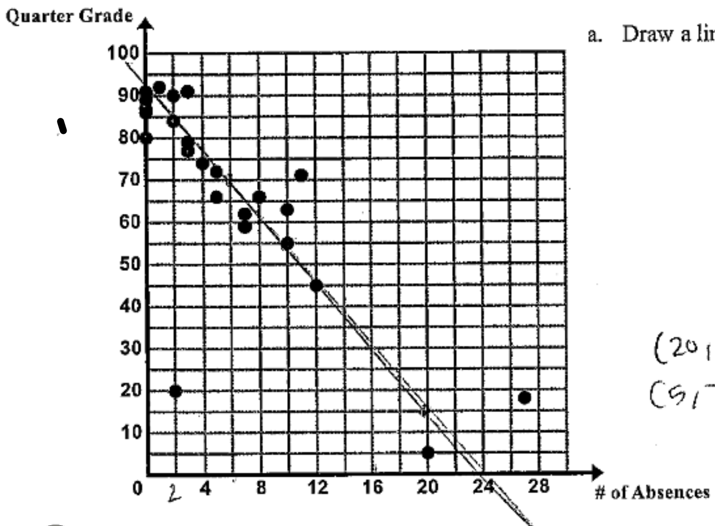
d. Write a prediction function for the data set. $y = \frac{1}{3}x + 0$

e. Use your prediction function to find the value of y when $x = 12$ and when $x = 100$.

12: $y = 4$

100: $y = 33.\overline{3}$

3 The following scatter plot shows the final quarter grade in Ms. Ganchero's math class for students vs. the number of times they are absent.



a. Draw a line of best fit on the scatter plot.

(20, 15)
(5, 70)

$\frac{70-15}{5-20} = \frac{55}{-15} = -\frac{11}{3}$

b. Write a prediction function for the line of best fit you drew.

$y = -\frac{11}{3}x + 90$

c. Explain the meaning of the slope and y-intercept in the context.

y-int: 0 absences ⇒ 90 on quarter grade

slope: your grade grade drops 11 for every 3 absences

d. Use your prediction function to predict the final grade of a student who is absent 16 times.

$y = -\frac{11}{3}(16) + 90 = 31.3$ Quarter Grade

e. Use your prediction function to predict how many times a student is absent who receives a final grade of 5 in the class.

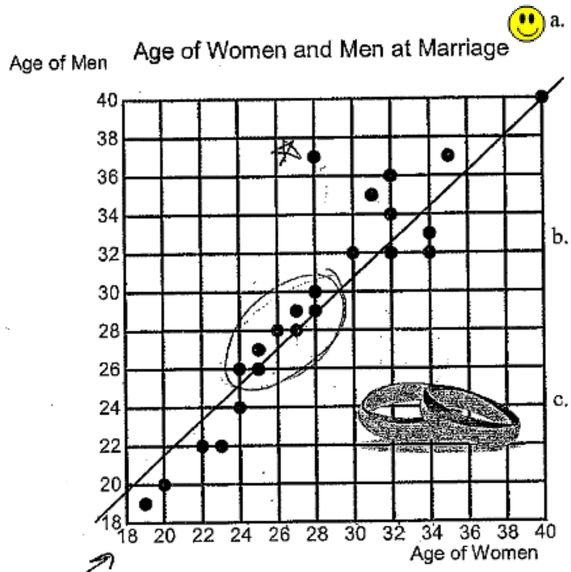
$5 = -\frac{11}{3}x + 90$

$-90 \qquad -90$

$\left(\frac{-3}{11}\right) -85 = \cancel{-\frac{11}{3}}x \cancel{\left(\frac{-3}{11}\right)}$

$x \approx 23.2$ so 23 times

4 Bethany is interested in the relationship between the age of when men and women get married. She surveys 24 couples and asks them the age in which they got married for the first time. A scatter plot of her data is below.



a. Describe the association between the two variables. Circle any clusters in the data. Put a star by any points that appear to be outliers.

pos, linear
cluster 24-28 women
26-30 men

b. Provide an explanation for any clusters of data or outliers.

mid-late 20's is a popular age
outlier - older man, younger wife

c. Draw a line of best fit on the scatter plot.

not 0....

d. Write a prediction function for the line of best fit you drew.

(27, 28)
(28, 29)
$\frac{1}{1}$ $y = x$ ~~+20~~
$y = x$

e. Use your prediction function to predict the age of a man when he gets married if the woman that he marries is 38.

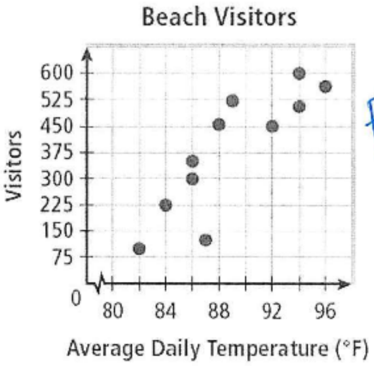
38 = 38

38 yrs old

8WB6 - 52

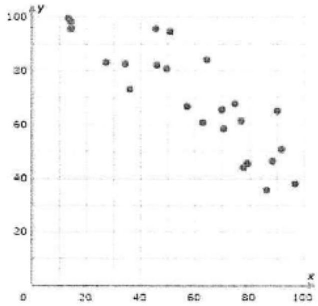
Describe the correlation of each scatter plot and scenario.

5.



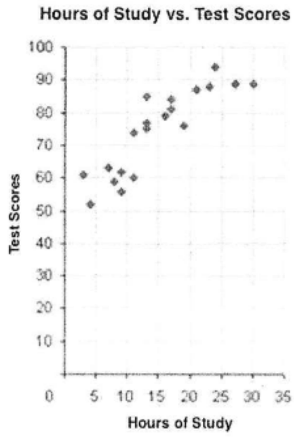
pos

6.



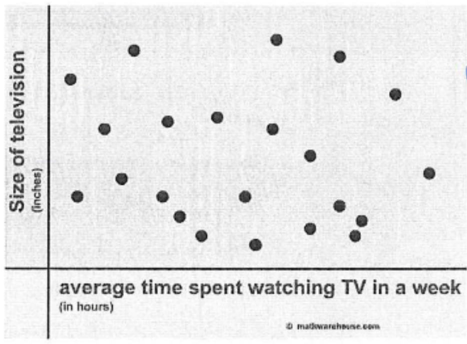
neg

7.



pos

8.



none

9. Temperature and number of popsicles sold.

pos

10. Number of books read and speeding tickets received.

none

11. Hours spent watching TV and test scores.

neg

12. Number of times you complain each day and number of friends you have.

neg (none if your friends REALLY love you).

13. Hours spent working and money earned.

pos

14. Number of hours spent practicing free throws and free throws made in a game.

pos

15. Hours spent on phone and number of shoes owned.

none

6.4B # 18-21 due tomorrow

Step 1 - put x's in L1 and y's in L2

Step 2 - look at scatter plot to determine if linear or exponential regression best fits the data

Step 3 - Do linear or exponential regression to get line (or curve) of best fit equation

Step 4 - look at correlation coefficient to determine if line (or curve) is a good fit (r for linear, r^2 for exponential)

20 minutes to work on hw from yesterday

When finished - grab rockin' residuals ws

8.8 Rockin' the Residuals

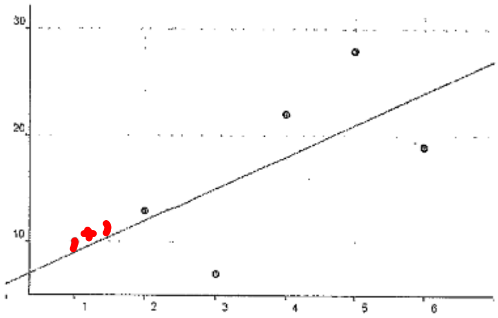
A Solidify Understanding Task

The correlation coefficient is not the only tool that statisticians use to analyze whether or not a line is a good model for the data. They also consider the residuals, which is to look at the difference between the observed value (the data) and the predicted value (the y-value on the regression line). This sounds a little complicated, but it's not really. The residuals are just a way of thinking about how far away the actual data is from the regression line.

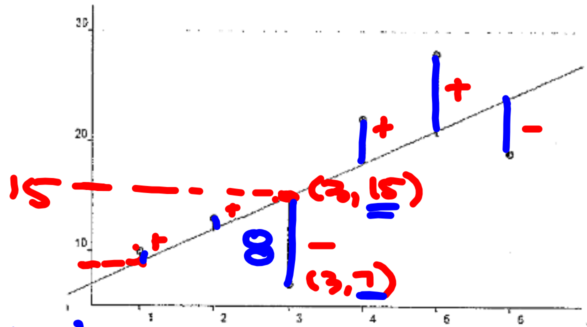
Start with some data:

x	1	2	3	4	5	6
y	10	13	7	22	28	19

Create a scatter plot and graph the regression line. In, this case the line is $y = 3x + 6$.



Draw a line from each data point to the regression line, like the segments drawn from each point below.

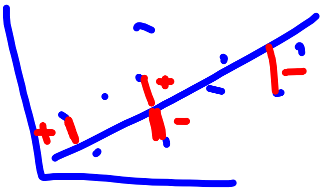


$y=\frac{1}{2}x+4$

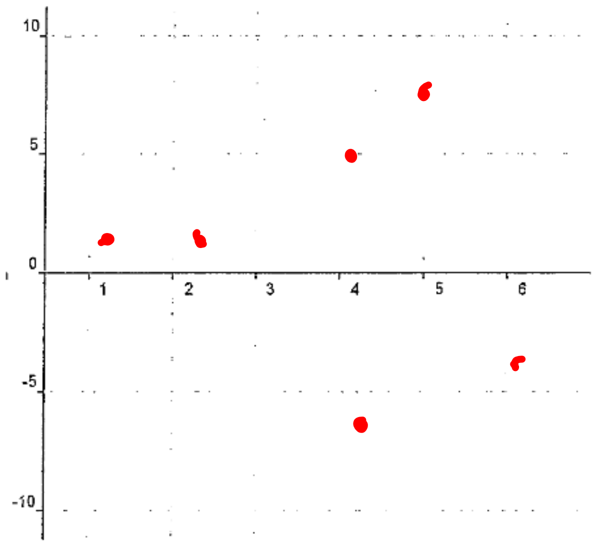
1. The residuals are the lengths of the segments. How can you calculate the length of each segment to get the residuals?

Distance between observed data & prediction line or distance formula $y = \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2}$

2. Generally, if the data point is above the regression line the residual is positive, if the data point is below the line, the residual is negative. Knowing this, use your plan from #1 to create a table of residual values using each data point.



3. Statisticians like to look at graphs of the residuals to judge their regression lines. So, you get your chance to do it. Graph the residuals here.



Now, that you have constructed a residual plot, think about what the residuals mean and answer the following questions.

4. If a residual is large and negative, what does it mean?

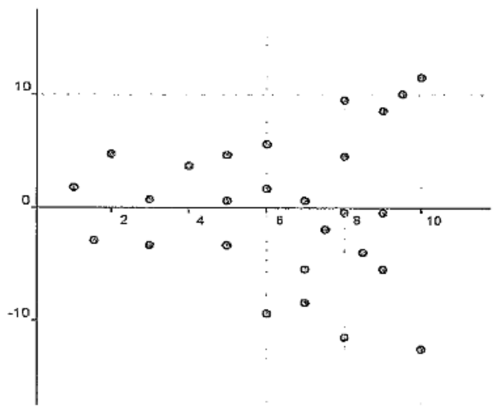
5. What does it mean if a residual is equal to 0?

6. If someone told you that they estimated a line of best fit for a set of data points and all of the residuals were positive, what would you say?

7. If the correlation coefficient for a data set is equal to 1, what will the residual plot look like?

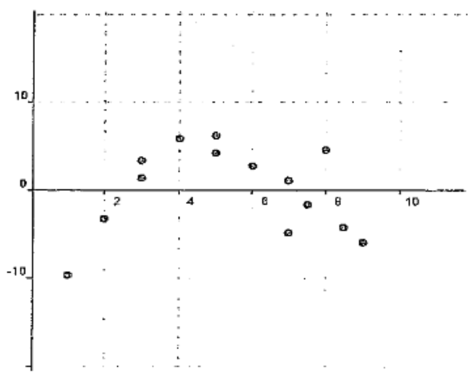
Statisticians use residual plots to see if there are patterns in the data that are not predicted by their model. What patterns can you identify in the following residual plots that might indicate that the regression line is not a good model for the data? Based on the residual plot are there any points that may be considered outliers?

8.

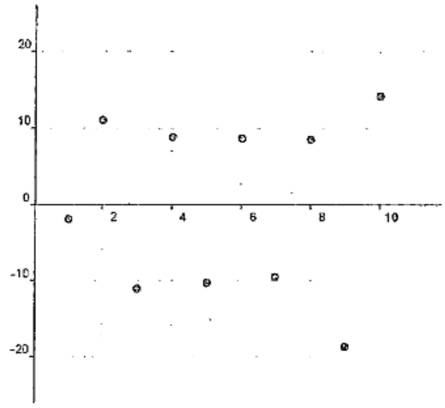


45

9.



10.



11.