

Bell Ringer

Section 6.4

For problems 1 and 2, the table below shows the amount of time that 8 basketball players spent practicing versus the number of baskets they made in the game.

Practice Time vs. Baskets Made

Practice Time (h)	Baskets Made
5	4
6	12
2	3
3	6
9	8
11	10
8	3
7	5

1. Make a scatter plot of the data.

2. What type of relationship does the scatter plot show?

Review.

3. The table represents a line. Write the equation of the line and a line perpendicular to it that passes through (3, 8).

x	y
3	-8
4	-10
5	-12
6	-14

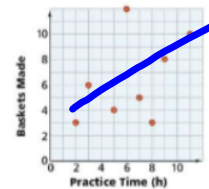
Solutions

Section 6.4

For problems 1 and 2, the table below shows the amount of time that 8 basketball players spent practicing versus the number of baskets they made in the game.

Practice Time vs. Baskets Made

Practice Time (h)	Baskets Made
5	4
6	12
2	3
3	6
9	8
11	10
8	3
7	5



1. Make a scatter plot of the data.

2. What type of relationship does the scatter plot show?

Positive Correlation

Review.

3. The table represents a line. Write the equation of the line and a line perpendicular to it that passes through (3, 8).

table $y = -2x - 2$ perpendicular $y = \frac{1}{2}x + \frac{13}{2}$

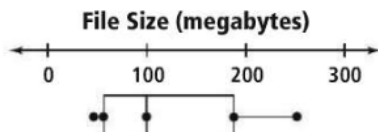
$$y = \frac{1}{2}x + b$$

x	y
3	-8
4	-10
5	-12
6	-14

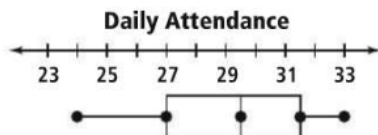
2 - 2
 $y = -\frac{2}{1}x$

correct 6.3 #s 8-16, 19-22

8. 48, 54, 100, 188, 256



9. 24, 27, 29.5, 31.5, 33



10. Class B

11. the box

12. 88

13. 75%; the third quartile is the value that divides the data so that about 75% of the data lies below and about 25% of the data lies above.

14. No; the test is scored on point values from 0 to 100, whereas the percentile rank tells you how you did in reference to the rest of the group.

15. $90 \leq x < 94$

16. The range gives the difference between the greatest and least values, while the interquartile range gives the difference between the third and first quartiles.

17. $0 < h < 73.5$ 18. Check students' work.

19. No; it could also be equal to the maximum value, which could happen if the top quarter of the scores all have the same value.

20. a. Device 1; the range is smaller.

- b. Device 2; Device 1; Device 2 has 25% of its packages below 17 oz, while Device 1 has more than that. Device 1 has 25% of its packages greater than 17.2 oz, while Device 2 has about 50% of its packages greater.

21. You cannot determine all of the data values by looking at the plot, so you cannot determine the mean or the mode. You can only find the median, which is the value at the line in the box.

22. 91

6.4A Trendlines ws due tomorrow
questions?

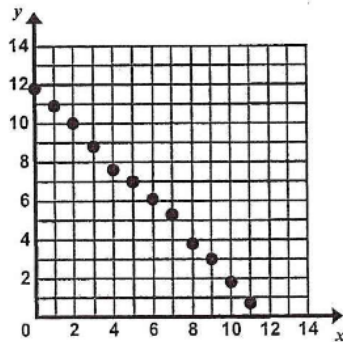
Skip #1 and #4

6.4A-Trendlines

Name _____ Date _____ Period _____

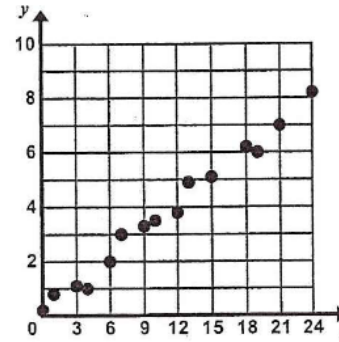
Directions: In #1 and 2, observe the data sets and take note of any associations you see, draw a line of best fit, write a prediction function, and use your function to predict the value of y when $x = 12$ and when $x = 100$.

1.



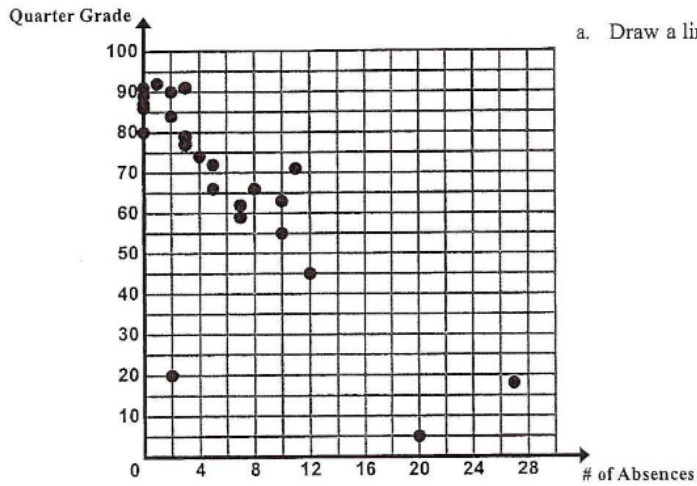
- Observations:
- Using a ruler, draw a line of best fit through the data points that captures the general trend of the data.
- Estimate the slope and y -intercept of your line.
 $m \approx$ _____ $b \approx$ _____
- Write a prediction function for the data set.
- Use your prediction function to find the value of y when $x = 12$ and when $x = 100$.

2.



- Observations:
- Using a ruler, draw a line of best fit through the data points that captures the general trend of the data.
- Estimate the slope and y -intercept of your line.
 $m \approx$ _____ $b \approx$ _____
- Write a prediction function for the data set.
- Use your prediction function to find the value of y when $x = 12$ and when $x = 100$.

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.

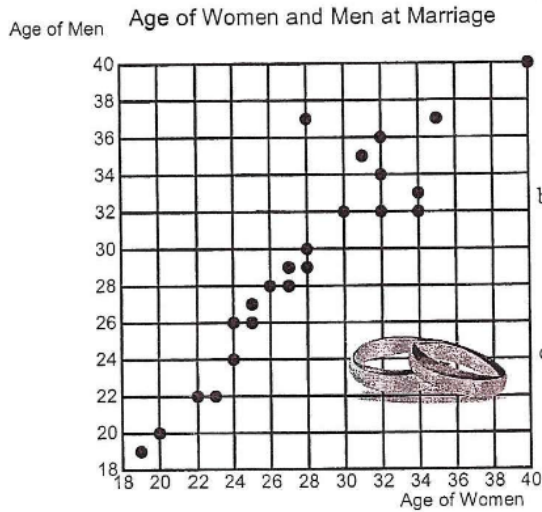
b. Write a prediction function for the line of best fit you drew.

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

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

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

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.

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

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

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

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

pg 373



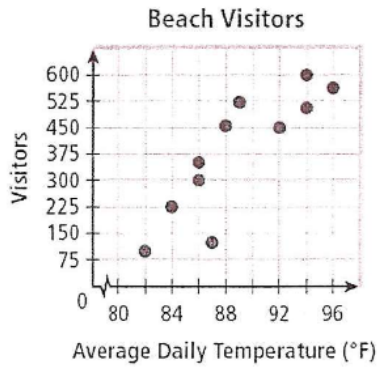
6-4 Scatter Plots and Trend Lines

Objectives To write an equation of a trend line and of a line of best fit
To use a trend line and a line of best fit to make predictions

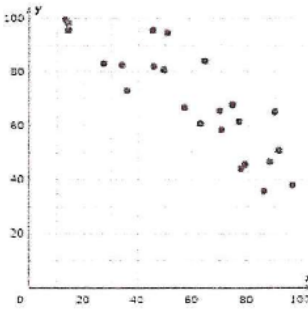
p373

Describe the correlation of each scatter plot and scenario.

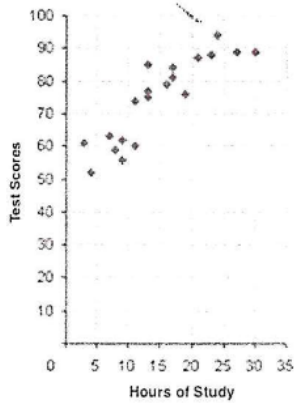
5.



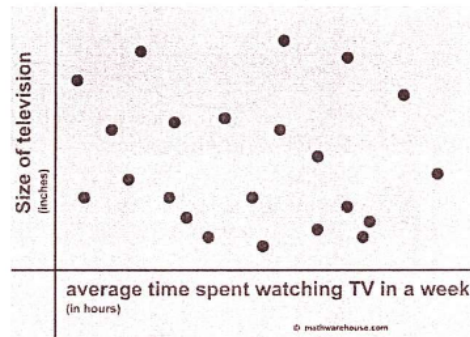
6.



7. Hours of Study vs. Test Scores



8.



9. Temperature and number of popsicles sold.

10. Number of books read and speeding tickets received.

11. Hours spent watching TV and test scores.

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

13. Hours spent working and money earned.

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

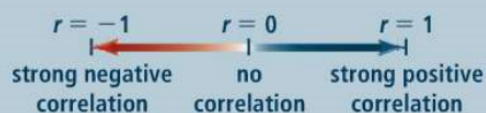
15. Hours spent on phone and number of shoes owned.

Trendlines vs Line of Best Fit...	
Approximation	Calculate it
Eye ball it	Exact

pg 378

The trend line that shows the relationship between two sets of data **most accurately** is called the **line of best fit**. A graphing calculator computes the equation of the line of best fit using a method called linear regression.

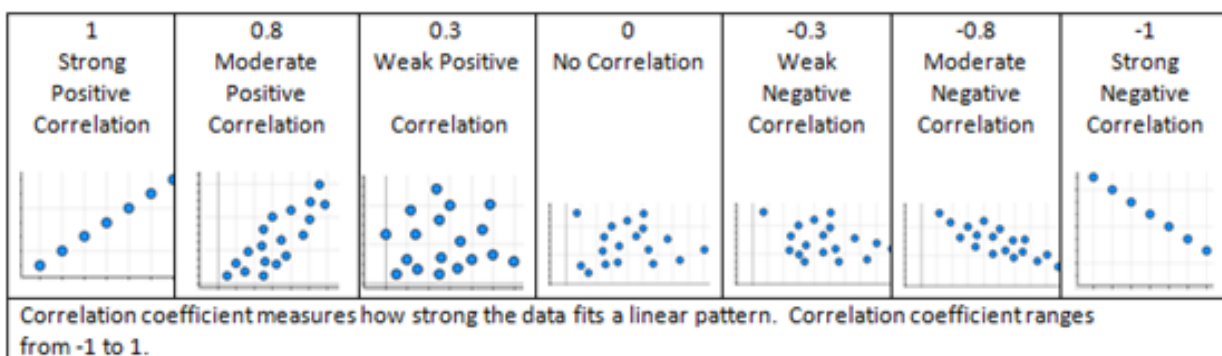
The graphing calculator also gives you the **correlation coefficient** r , a number from -1 to 1 , that tells you how closely the equation models the data.



The nearer r is to 1 or -1 , the more closely the data cluster around the line of best fit. If r is near 1 , the data lie close to a line of best fit with positive slope. If r is near -1 , the data lie close to a line of best fit with negative slope.

p378

Correlation coefficient, r



EQUATIONS

Linear: $y = mx + b$
 $y = ax + b$

Exponential: $y = a \cdot b^x$

You can get a line of BEST fit by using the calculator...

1

MODE
Stat Diagnostics..... ON
Enter

2

y = delete everything

not in book


Problem 3 Finding the Line of Best Fit

College Tuition Use a graphing calculator to find the equation of the line of best fit for the data at the right. What is the correlation coefficient to three decimal places? Predict the cost of attending in the 2012–2013 academic year.

Linear Regression:

$$y = 409.4x - 865446.7$$

Correlation Coefficient?

$$r = .99$$

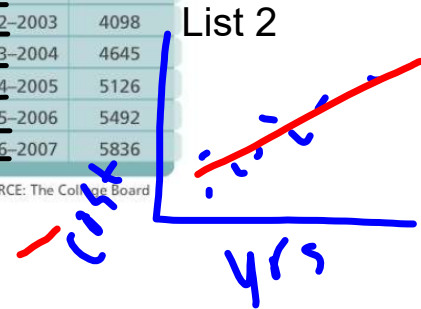
Exponential Regression:

Correlation Coefficient?

Average Tuition and Fees at Public 4-Year Colleges

Academic Year	Cost (\$)
2000–2001	3508
2001–2002	3766
2002–2003	4098
2003–2004	4645
2004–2005	5126
2005–2006	5492
2006–2007	5836

SOURCE: The College Board



solution...

**Problem 3** Finding the Line of Best Fit

Round to the nearest hundredth. The equation of the line of best fit is $y = 409.43x - 815,446.71$. The correlation coefficient is about 0.996.

Use the equation of the line of best fit.

$$y = 409.43x - 815,446.71$$

Substitute 2012 for x .

$$y = 409.43(2012) - 815,446.71$$

Simplify. Round to the nearest whole number.

$$y \approx 8326$$

The cost of attending a four-year public college in the 2012–2013 academic year is predicted to be about \$8326.



The correlation coefficient of 0.996 is close to 1. So there is a strong positive correlation between the academic year and the cost of attending college.

got it pg 378

- Got It?** a. Use the data in Problem 3. Predict the cost of attending in the 2016-2017 academic year.

$$y = 409.43x - 815,446.71$$

$$409.4(2016) - 815,446.7 = 9,964.17$$

- © b. **Reasoning** What does the slope of the line of best fit in Problem 3 tell you about the rate of change in the cost?

p378

not in book

**Problem 5 Analyzing a Nonlinear Trend**

Gaming The table shows the number of members of an online gaming forum during the first several months after the forum was formed. Use a graphing calculator to make a scatter plot for the data. Which is a suitable model for the data, a linear model or an exponential model? Predict the number of members after 10 months.

Online Forum Membership

Month, x	Number of Members, y
1	20
2	41
3	67
4	118
5	209
6	361
7	632

Linear Regression:

Correlation Coefficient?

Exponential Regression:

Correlation Coefficient?

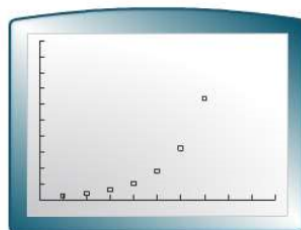
solution

**Problem 5** Analyzing a Nonlinear Trend

Gaming The table shows the number of members of an online gaming forum during the first several months after the forum was formed. Use a graphing calculator to make a scatter plot for the data. Which is a suitable model for the data, a linear model or an exponential model? Predict the number of members after 10 months.

Online Forum Membership

Month, x	Number of Members, y
1	20
2	41
3	67
4	118
5	209
6	361
7	632



The scatter plot shows a nonlinear trend. The membership data appear to be exponential.

The membership data is nonlinear. An exponential model of the data is $y = 12.21(1.76)^x$.


Substitute 10 for x . $y = 12.21(1.76)^{10}$

Simplify. $y \approx 3482$

There will be about 3482 members after 10 months.

Got It? The table shows the value of Denise's car each year since she purchased it. Use a graphing calculator to make a scatter plot for the data. Which is a suitable model for the data, a linear model or an exponential model? Write the model and predict the value of Denise's car after 8 years.

How can a scatter plot show whether a linear or exponential model is better?



Value of Denise's Car					
Age (years), x	1	2	3	4	5
Value (thousands of dollars), y	28.6	18.5	14.1	10.2	8.9

Linear Regression:

$$y = 4.77x + 30.3$$

Correlation Coefficient?

$$r = -.94$$

Exponential Regression:

$$y = 35.3(.74)^x \approx 3.2$$

Correlation Coefficient?

$$r^2 = .97$$

p380

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.

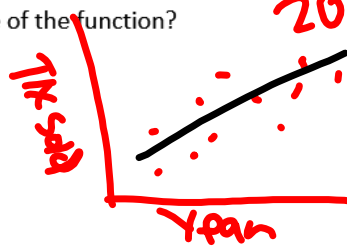
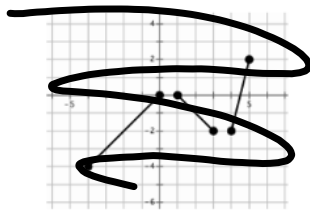
Movie Tickets Sold in U.S. by Year										
Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Tickets Sold (millions)	1289	1311	1340	1339	1406	1421	1470	1415	1472	1470

Source: Motion Picture Association of America

LoF B.F. $y = 21.4x - 41,557.3$
 $r = .94$ $r = .32$
 2014: 1,542 tickets

Review.

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



Correlation = Causation...???

Correlation is a statistical measure (expressed as a number) that describes the size and direction of a relationship between two or more variables. A correlation between variables, however, does not automatically mean that the change in one variable is the cause of the change in the values of the other variable.

Causation indicates that one event is the result of the occurrence of the other event; i.e. there is a causal relationship between the two events. This is also referred to as cause and effect.

Theoretically, the difference between the two types of relationships are easy to identify — an action or occurrence can cause another (e.g. smoking causes an increase in the risk of developing lung cancer), or it can correlate with another (e.g. smoking is correlated with alcoholism, but it does not cause alcoholism). In practice, however, it remains difficult to clearly establish cause and effect, compared with establishing correlation.

-Australian Bureau of Statistics

**Problem 4** Identifying Whether Relationships Are Causal

In the following situations, is there likely to be a correlation? If so, does the correlation reflect a causal relationship? Explain.

A the number of loaves of bread baked and the amount of flour used

B the number of mailboxes and the number of firefighters in a city

**Problem 4** Identifying Whether Relationships Are Causal

In the following situations, is there likely to be a correlation? If so, does the correlation reflect a causal relationship? Explain.

A the number of loaves of bread baked and the amount of flour used

As the number of loaves of bread baked increases, the amount of flour used increases.

There is a positive correlation and also a causal relationship.

**Think**

Remember, causal relationships always have a correlation. However, two data sets that have a correlation may not have a causal relationship.

B the number of mailboxes and the number of firefighters in a city

There is likely to be a positive correlation because both the number of mailboxes and the number of firefighters tend to increase as the population of a city increases.

However, installing more mailboxes will not cause the number of firefighters to increase, so there is no causal relationship.

got it pg 379

Got It? In the following situations, is there likely to be a correlation? If so, does the correlation reflect a causal relationship? Explain.

a. the cost of a family's vacation and the size of their house

b. the time spent exercising and the number of Calories burned

p379

hw 6.4B Lines of Best fit ws

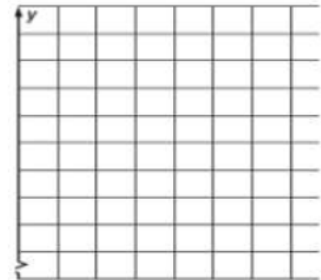
due Friday

Name _____ Hour _____ 6.4B Lines of Best Fit

1. Explain the difference between a trend line and a line of best fit.
2. The table shows number of hours spent studying for a science test and final test score.

Study Hours	3	2	5	1	0	4	3
Grade	84	77	92	70	60	90	75

- Draw a scatter plot of the data and a line of fit (trend line).
- Write a **linear function** for the line of BEST fit and interpret The correlation coefficient. Is the line a good fit? Explain.
- Predict the grade of a student who studied 6 hours.
- Is this interpolation or extrapolation?



3. The table shows the average and maximum longevity of various animals in captivity.

		Longevity (years)							
L. x	Avg.	12	25	15	8	35	40	41	20
L. y	Max.	47	50	40	20	70	77	61	54

- Draw a scatter plot and determine what relationship, if any, exists in the data.
- Draw a line of fit (trend line) for the scatterplot, then use your calculator to write the linear function of the line of BEST fit and interpret the correlation coefficient.

$y = 1.22x + 22.45 \quad r = .88$

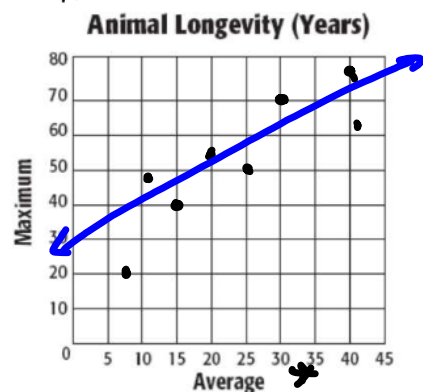
- Is the line a good fit? Explain.

Yes, because r is close to 1

- Predict the maximum longevity for an animal with an average longevity of 33 years.

$y = 1.22(33) + 22.45 \approx 62.7 \text{ yrs}$

- Is this interpolation or extrapolation?



4. A herd of caribou moved to a small remote island where they had no predators. Data on the population of the herd was collected for 6 years.

Time (years)	0	1	2	3	4	5	6
Population	24	35	51	74	104	151	225



a. Write the equation for the line of best fit.
(linear regression)

$$y = 31.7x - .29$$

b. $r = .95$

c. Write the equation for the curve of best fit.
(exponential regression)

$$y = 24.1(1.45)^x$$

d. $r^2 = .9997$

e. Is the linear or exponential regression a better fit for the data?

Linear

Exponential

Closer to 1

f. Using the equation that best fits the data, predict how many caribou there will be in 9 years.

$$y = 24.1(1.45)^9$$

g. Is this interpolation or extrapolation?

683 caribou

5. Use the data in the table to answer the questions.

x	1	1	2	5	4	4	3	6
y	9	3	8	4	3	2	6	8

a. Write the equation for the line of best fit.
(linear regression)

b. $r = \underline{\hspace{2cm}}$

c. Write the equation for the curve of best fit.
(exponential regression)

d. $r^2 = \underline{\hspace{2cm}}$

e. Is the linear or exponential regression a better fit for the data?

Linear

Exponential

f. Using the equation that best fits the data, predict y when x is 3.

g. Is this interpolation or extrapolation?