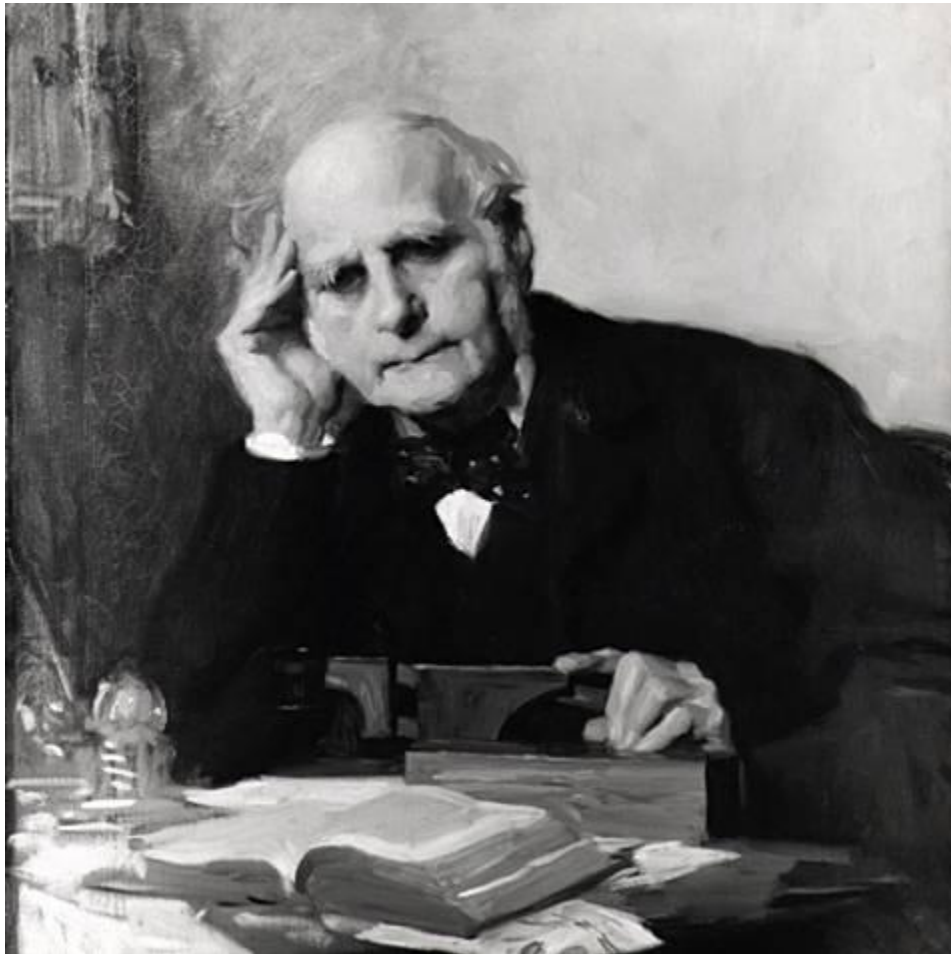


Linear Functions, Models, and Regression in the Real, Real World

*Scott Adamson
Chandler-Gilbert Community College*

**Sir Francis Galton
(1822-1911)**

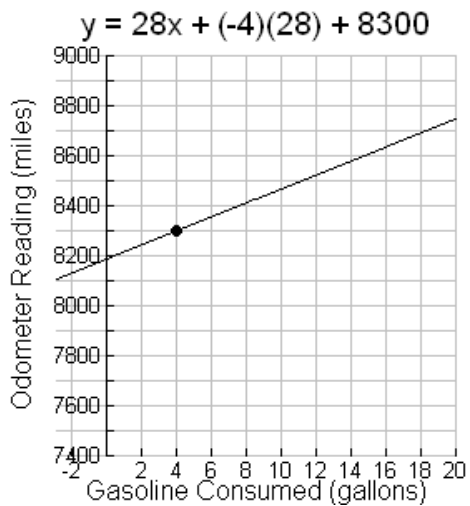


Do You REALLY Understand Linear Functions?

Part I

You are traveling in your car down the freeway with the cruise control set. You observe your odometer reading to be 8300 miles and that your car has consumed 4 gallons of gasoline. Your car gets 28 miles per gallon. Create a linear function that models the odometer reading as a function of gasoline consumed. Write a detailed explanation of what this linear function represents in the context of the situation.

Part II



1. In Part I, you were asked to define a function so that its graph passes through the point (4, 8300) with a rate of change of 28. This figure shows one function that does this.

- a) What does 28 stand for in this figure?

- b) What does -4 stand for in this figure?

- c) What does $(-4)28$ stand for in this figure?

- d) What does $(-4)28+8300$ stand for in this figure?

- e) What does $28x + (-4)28+8300$ stand for in this figure?

Linear Functions, Models, and Regression in the Real, Real World
2009 CMC³

2. Do the following for each of a-e. Check your answers on your graphing calculator.
Define a function in the form $y = ax + b$ so that its graph passes through

a) The point (5,2) with a rate of change of 7.

b) The point (-3, 4) with a rate of change of -2.

c) The point (2.73,-5.15) with a rate of change of 7.26

d) The point (-4.1,-6.8) with a rate of change of 8.6

e) The point (-2,-1) with a rate of change of -4.2

Linear Functions, Models, and Regression in the Real, Real World
2009 CMC³

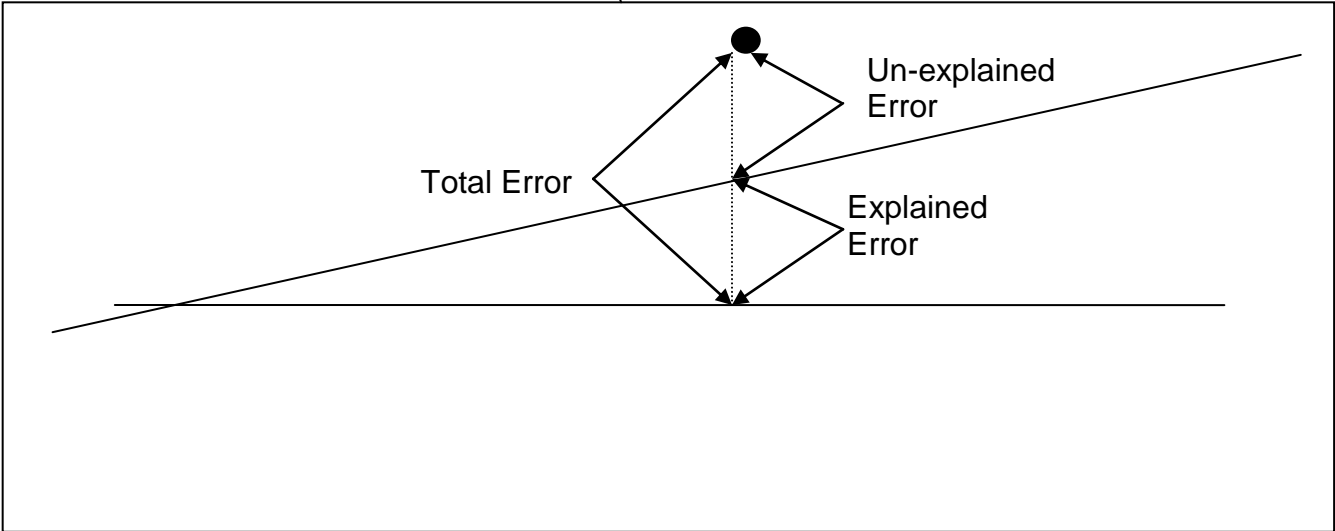
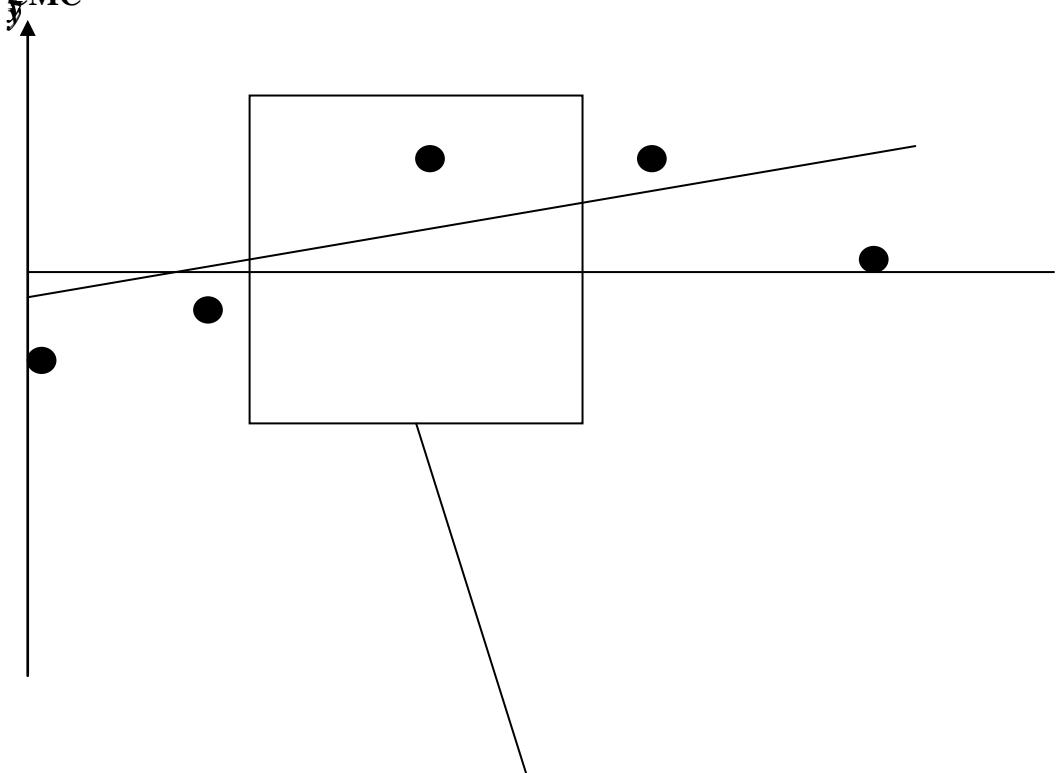
3. Generalize your previous work by writing a function definition for line that passes through the point (j, k) with a constant rate of change of h . Create a graphical representation that could be used to help explain this general function definition.

Linear Functions, Models, and Regression in the Real, Real World
 2009 CMC³

Year (t=0 is 1980)	Number of Cars (millions)	Mean	Distance from the Mean	Square the Distance from the Mean
0	161			
5	177			
10	193			
13	198			
16	210			
			SSTotal	

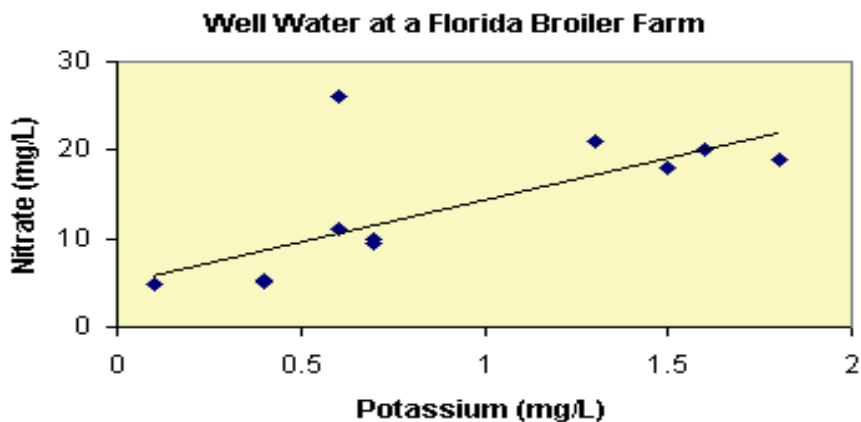
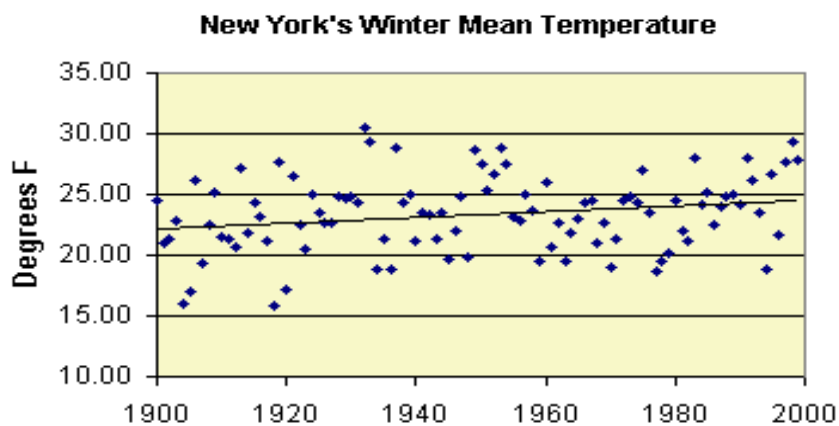
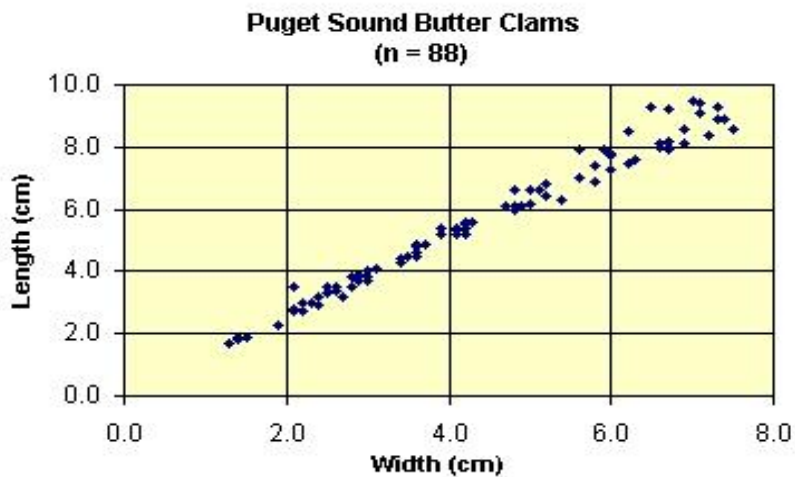
Year (t=0 is 1980)	Number of Cars (millions) reported by model	Mean	Distance from the Mean	Square of the distance from the Mean
0	161.54			
5	176.46			
10	191.38			
13	200.33			
16	209.29			
			SSRTotal	

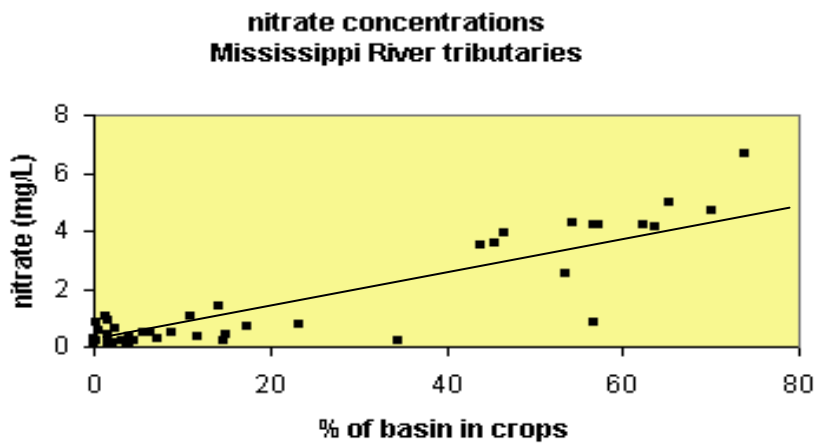
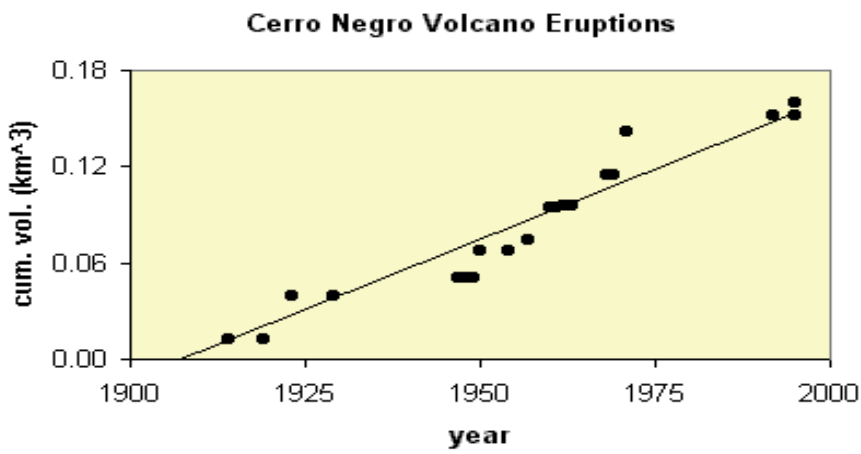
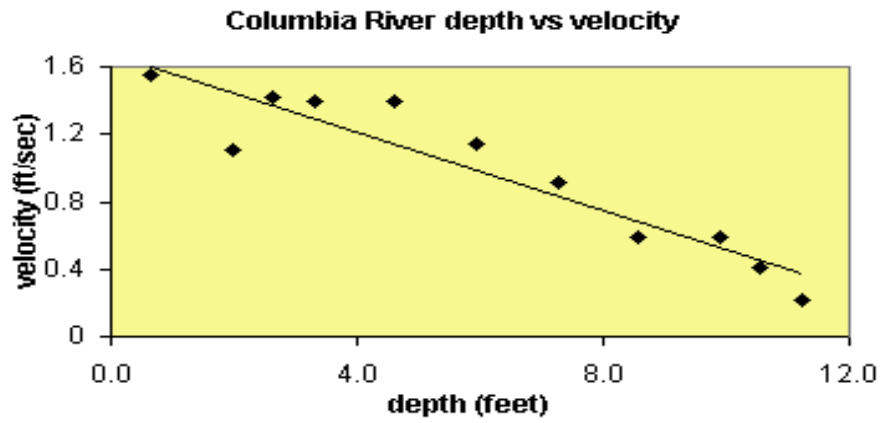
Year (t=0 is 1980)	Number of Cars (millions)	Number of Cars (millions) reported by model	Distance between actual and model	Square of the distance from the mean
0	161	161.54		
5	177	176.46		
10	193	191.38		
13	198	200.33		
16	210	209.29		
			SSETotal	



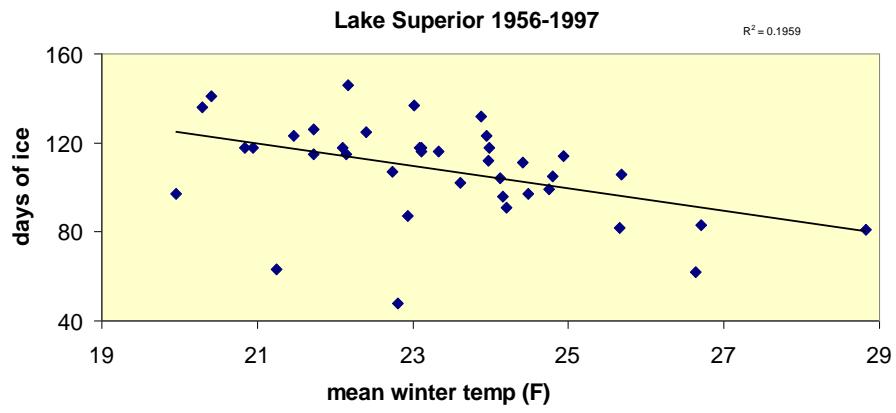
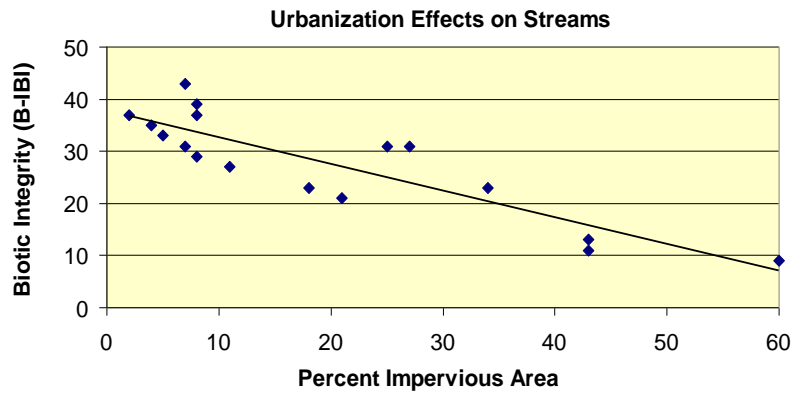
The Coefficient of Determination

For each of the following, state whether you think that the coefficient of determination, R^2 , is closer to 1, closer to 0.5, or closer to 0. Explain your thinking (Source: www.seattlecentral.edu/qelp/Data.html)





Linear Functions, Models, and Regression in the Real, Real World
2009 CMC³



The Hand Squeeze

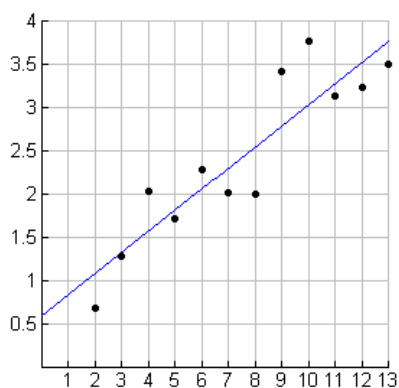
Set Up

Start with two people standing face-to-face holding hands (like at a wedding). A third student can be the time keeper. When the time keeper says go, the designated squeeze starter squeezes one of the hands of the second person. Upon feeling the squeeze, this second person squeezes back with the other hand. The squeeze starter shouts “Stop!” when the squeeze is felt and the time keeper stops the stopwatch. Record the time. Continue this procedure by adding one person to the circle each iteration.

Sample Data

The following data were collected in an Intermediate Algebra class in the Spring of 2004.

People	2	3	4	5	6	7	8	9	10	11	12	13
Time (seconds)	0.69	1.29	2.04	1.72	2.29	2.02	2.00	3.41	3.76	3.13	3.24	3.50



Linear Regression ($ax+b$)
 $\text{regEQ}(x) = .243462x + .598205$
 $r = .903828$
 $r^2 = .816905$

Suggested Questions

1. Determine a linear function that models the data.
2. Describe the meaning of the constant rate of change in this situation.
3. Describe the meaning of the vertical intercept in this situation.
4. Using the linear model, determine the time it would take for the squeeze to travel from beginning to end if there were 30 people involved.
5. If the time keeper recorded a time of 10 seconds for the squeeze to travel from beginning to end, how many people were involved according to the linear model you found?
6. Explain the meaning of the coefficient of determination.
7. Explain the meaning of the correlation coefficient.
8. Since a linear model was used in this situation, we would say that the rate of change is constant. Explain what constant rate of change means in the context of this situation.
9. Suppose each person was precisely 0.20 seconds slower than what was recorded. How could you adjust the linear model found in #1 and write the equation of the linear function that models this new data set?

Linear Functions, Models, and Regression in the Real, Real World

2009 CMC³

10. Compare the slope of the linear function found in #1 with the slope of the linear function in #9. What do you notice? Why is this the case?

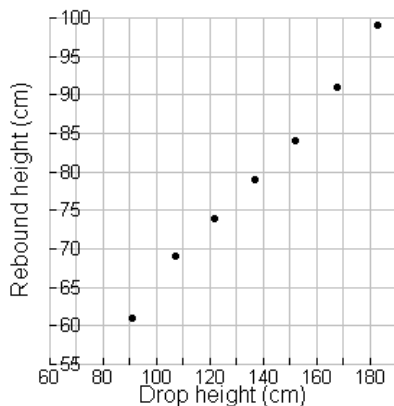
That's the Way the Ball Bounces

Set Up

In this activity, students measure the rebound height of different types of balls when they are dropped from different initial heights. Students will tape cloth tape measures to the wall and drop the ball from several different heights, carefully measuring the rebound height. It is often helpful to collect at least 3 rebound heights for each initial height, taking the average rebound height as the data point. It is interesting to give each group a different type of ball so that comparisons can be made. Golf balls, racquetball balls, tennis balls, ping pong balls, etc. all work well.

Sample Data – Ping Pong Ball

Initial Height (cm)	91	107	122	137	152	168	183
Rebound Height (cm)	61	69	74	79	84	91	99



Linear Regression ($ax+b$)
 $\text{regEQ}(x) = .392707x + 25.7144$
 $a = .392707$
 $b = 25.7144$
 $r^2 = .993423$

Suggested Questions

1. Determine a linear function that models the data.
2. Describe the meaning of the constant rate of change in this situation.
3. Describe the meaning of the vertical intercept in this situation.
4. Using the linear model, determine the rebound height for a ball dropped from an initial height of 200 cm.
5. If the rebound height recorded was 55 cm, what was the initial height from which the ball was dropped?
6. Explain the meaning of the coefficient of determination.
7. Explain the meaning of the correlation coefficient.
8. Since a linear model was used in this situation, we would say that the rate of change is constant. Explain what constant rate of change means in the context of this situation.
9. Suppose rebound height was precisely 2.5 cm less than what was recorded. How could you adjust the linear model found in #1 and write the equation of the linear function that models this new data set?

Linear Functions, Models, and Regression in the Real, Real World

2009 CMC³

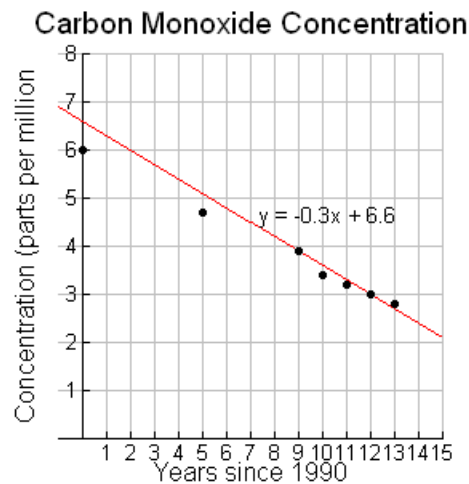
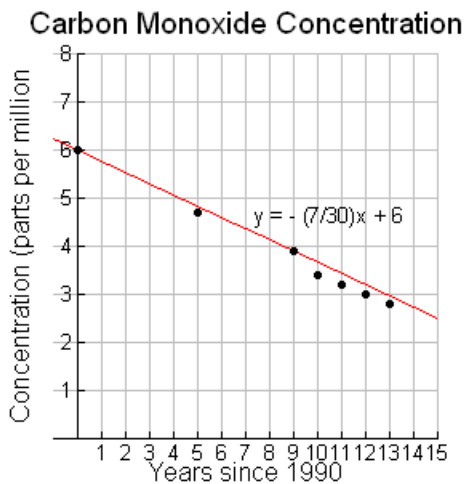
10. Compare the slope of the linear function found in #1 with the slope of the linear function in #9. What do you notice? Why is this the case?

Carbon Monoxide Concentration

The amount of carbon monoxide pollutant concentration (in parts per million) for particular years is given in the table (Source: Statistical Abstract of the United States, 2006, Table 359).

t Year (Since 1990)	P Carbon Monoxide Pollutant Concentration (parts per million)
0	6
5	4.7
9	3.9
10	3.4
11	3.2
12	3
13	2.8

1. Determine which linear function best models these data as determined by the total sum of squares.



2. Use a calculator to determine a linear regression model for these data. Explain what the coefficient of determination and the correlation coefficient tells you about the strength of fit of this model.
3. Calculate the total sum of the squares of the residuals of the regression model to show that it is less than those found in #1.
4. Interpret the meaning of the slope and the vertical intercept of the model in terms of carbon monoxide pollutant concentration, P , and years since 1990, t .
5. Use the model to predict the carbon monoxide pollutant concentration in 2008. Would this be an example of interpolation or extrapolation? Why?

Sosa and McGwire

The 1998 Major League Baseball season was, as they say, “One for the record books!” Imagine that the date is June 8th, 1998. Sammy Sosa of the Chicago Cubs has just hit his 20th homerun of the season. Mark McGwire of the St. Louis Cardinals has just hit his 29th homerun of the season. Sports fans are beginning to say that the all-time homerun record set by Roger Maris in 1961 is going to fall.

Mark McGwire	
Game (g)	Total number of home runs hit (h)
1	1
2	2
3	3
4	4
13	7
16	8
19	9
23	10
27	11
28	12
34	13
36	14
38	15
40	16
42	17
43	19
44	20
46	21
47	23
48	24
49	25
52	26
53	27
59	28
62	29

Sammy Sosa	
Game (g)	Total number of home runs hit (h)
5	1
11	2
15	3
22	4
23	5
26	6
31	7
43	8
48	9
51	11
52	13
57	15
59	16
60	17
61	18
62	19
63	20

1. Make a scatter plot of the data points.
2. Calculate a linear model for each data set.
3. For each model, interpret the meaning of the slope and the vertical intercept in the context of the situation.
4. Determine the practical domain and range for each model. Explain.
5. For the data provided, who always had more home runs? How can you tell?
6. Determine the number of home runs, using your models, that each player will hit for the season (there are 162 games in a season).
7. According to your models, when will the number of homeruns for each player be 61? Discuss the accuracy of your results.

Linear Functions, Models, and Regression in the Real, Real World

2009 CMC³

8. Is the number of home runs hit by McGwire or Sosa increasing more rapidly? How can you tell?
9. Will the number of homeruns hit by each player ever be the same? If so, determine when this will occur.