

14.2 Choosing Among Linear, Quadratic, and Exponential Models

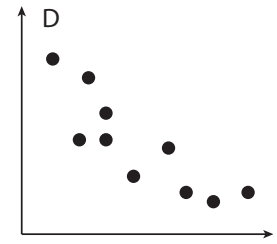
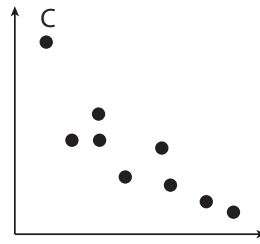
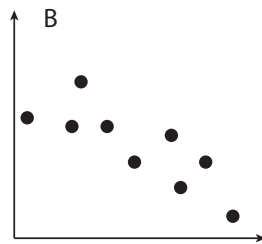
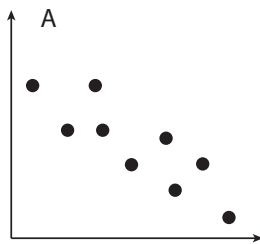


Resource Locker

Essential Question: How do you choose among, linear, quadratic, and exponential models for a given set of data?

Explore Developing Rules of Thumb for Visually Choosing a Model

When you work with data, you may not know whether a linear, quadratic, or exponential model will be a good fit. If the data lie along a curve that rises and then falls or falls and then rises, the data may be well-fitted by a quadratic model. But sometimes it may not be as clear. Consider the following scatter plots.



- A** Look at scatter plot A. Do you think a linear model will be appropriate? Explain your reasoning. If you think a linear model is appropriate, what do you know about the lead coefficient?

- B** Look at scatter plot B. What is different now that indicates that another kind of model might be appropriate? What characteristics would this model have?

- C** Look at scatter plot C. What about this plot indicates that yet another kind of model might be appropriate? What characteristics would this model have?

- D Look at scatter plot D. This plot is very similar to plot C, but what indicates that a different model would be appropriate? What characteristics would this model have?

Reflect

1. When can it be difficult to distinguish whether a quadratic or an exponential model is most appropriate?

2. Under what circumstances might it be difficult to tell exponential or quadratic data from linear data?

3. For data that do not lie tightly along a curve or line, what is different about the last data point that can make it potentially more misleading than other points?

Explain 1 Modeling with a Linear Function

As noted in the Explore, it is not always immediately clear what kind of model best represents a data set. With experience, your ability to recognize signs and reasons for choosing one model over another will increase.

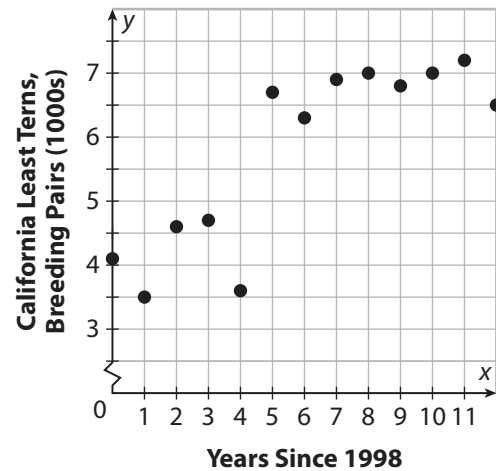
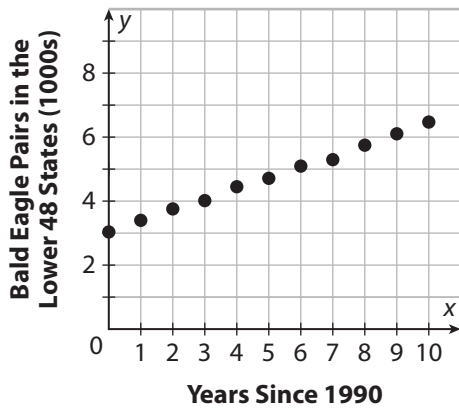
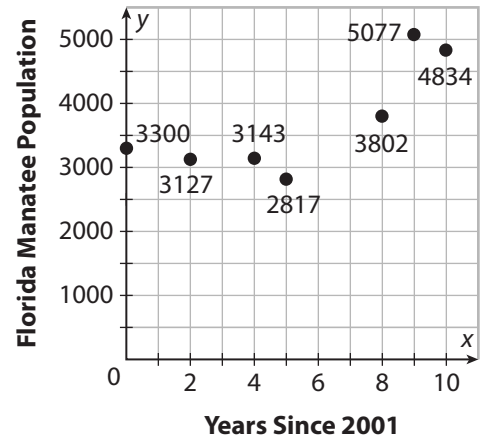
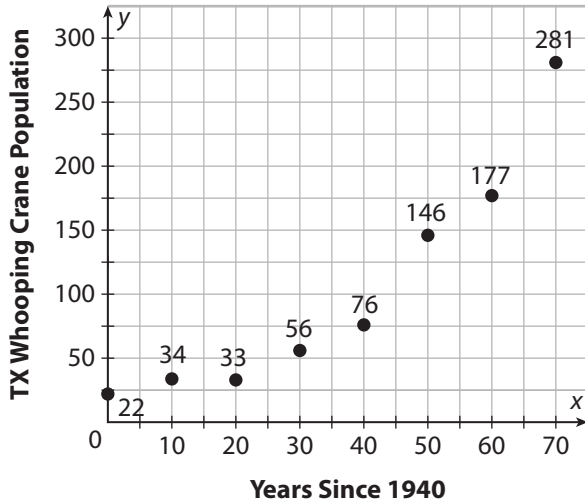


© Houghton Mifflin Harcourt Publishing Company. © L.M. Otero/AP Images

Example 1 Examine each scatter plot. Then complete the steps below.

- Step 1:** Choose the data set that appears to be best modeled by a linear function. Explain your choice, whether you think a linear model will be a close fit, and whether any other model might possibly be appropriate. What characteristics do you expect the linear model will have?
- Step 2:** Enter the data for your choice into your graphing calculator in two lists, and perform linear regression. Then give the model, defining your variables. What are the initial value and the rate of change of the model?
- Step 3:** Graph the model along with the scatter plot using your calculator, then assess how well the model appears to fit the data.

- A Wildlife Conservation** Data sets and scatter plots for populations over time of four endangered, threatened, or scarce species are shown.



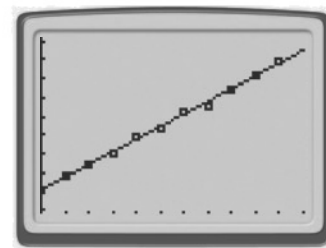
(1990, 3035)	(1994, 4449)	(1998, 5748)
(1991, 3399)	(1995, 4712)	(1999, 6104)
(1992, 3749)	(1996, 5094)	(2000, 6471)
(1993, 4015)	(1997, 5295)	

(1998, 4100)	(2003, 6700)	(2007, 6800)
(1999, 3500)	(2004, 6300)	(2008, 7000)
(2000, 4600)	(2005, 6900)	(2009, 7200)
(2001, 4700)	(2006, 7000)	(2010, 6500)
(2002, 3600)		

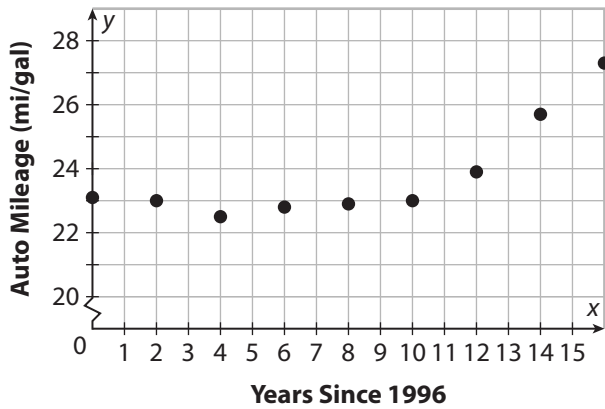
Step 1: The bald eagle population is clearly the one best modeled by a linear function, as the increase in the number of pairs is very steady, with no apparent curving or changes in the trend that might indicate a different model. The model will have a y -intercept of about 3 (in thousands) and will have a slope very close to 0.34, since the rate of change all along the graph remains close to the average rate of change from the first point to the last.

Step 2: A regression model is $y = 338.2x + 3042$ where x is the number of years after 1990 and y is the number of breeding pairs in thousands. The initial value is 3042, and the rate of change is about 338 pairs per year.

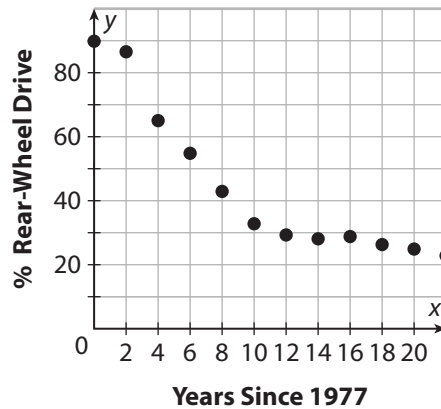
Step 3: The model is a very close fit to the data. It fits both the overall trend and the individual points very closely.



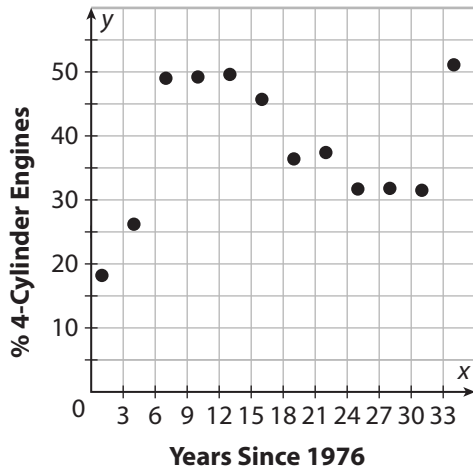
B Automobiles Data sets and scatter plots for various statistics about changes in automobiles of different model years are shown.



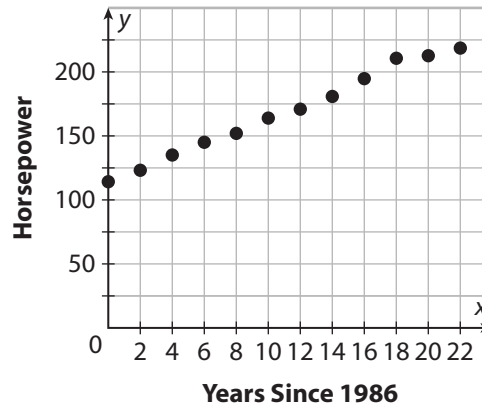
(1996, 23.1) (2002, 22.8) (2008, 23.9)
 (1998, 23.0) (2004, 22.9) (2010, 25.7)
 (2000, 22.5) (2006, 23.0) (2012, 27.3)



(1977, 89.8) (1985, 42.9) (1993, 28.8)
 (1979, 86.5) (1987, 32.8) (1995, 26.3)
 (1981, 65.0) (1989, 29.3) (1997, 24.9)
 (1983, 54.8) (1991, 28.1) (1999, 22.9)



(1976, 18.2) (1988, 49.6) (2000, 31.7)
 (1979, 26.2) (1991, 45.7) (2003, 31.8)
 (1982, 49.0) (1994, 36.4) (2006, 31.5)
 (1985, 49.2) (1997, 37.4) (2009, 51.1)



(1986, 114) (1994, 152) (2002, 195)
 (1988, 123) (1996, 164) (2004, 211)
 (1990, 135) (1998, 171) (2006, 213)
 (1992, 145) (2000, 181) (2008, 219)

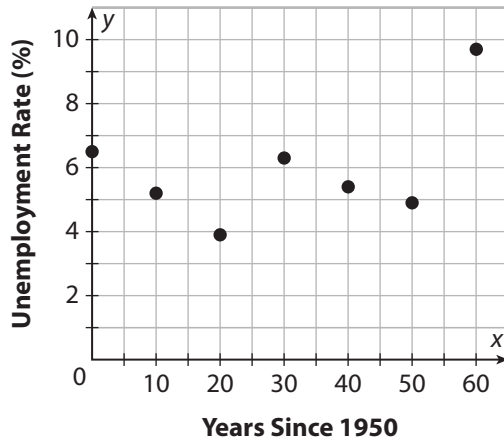
Step 1: The are the data best modeled by a linear function, as the increase in over time is very steady, with the amount of increase never varying too much from the trend. All of the other data sets show clear variation in the rate of change. The model will have a y -intercept of about and will have a positive slope that should be close to the overall average rate of change of , since the rate of change all along the graph remains close to the average rate of change from the first point to the last.

Step 2: A regression model is where x is the number of years after 1986 and y is the for the model year. The initial value is about 114 horsepower, and the rate of change is about .

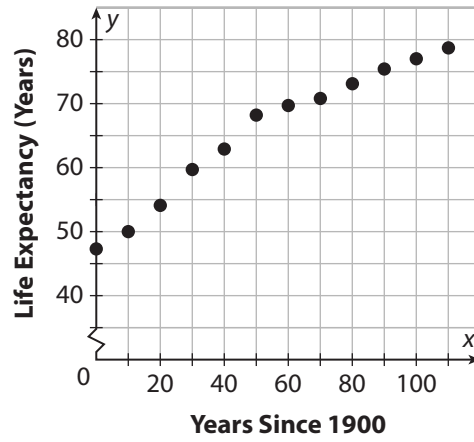
Step 3: The model fits both the overall trend and the individual points very closely.

Your Turn

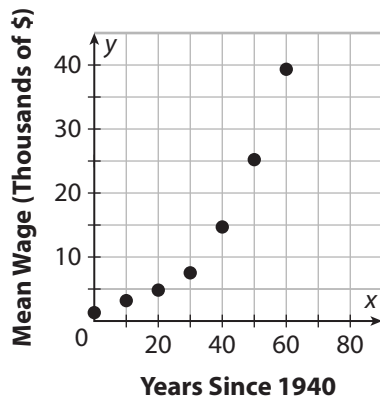
4. Demographics Data sets and scatter plots for various changes in the United States population over time are shown. Using these data, complete the three steps described at the beginning of Example 1. Also, tell whether you would expect the trend indicated by your model to continue for a time after the data shown, or whether you expect that it would soon change, and explain your answer.



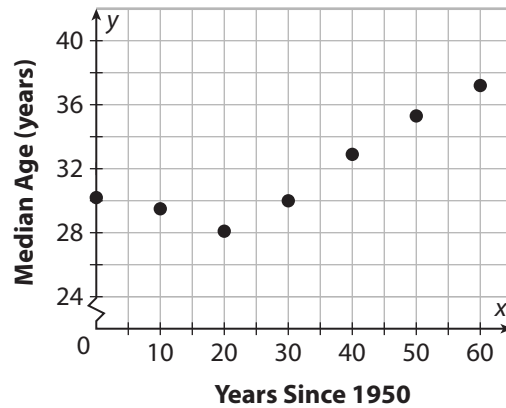
(1950, 6.5) (1980, 6.3) (2000, 4.9)
 (1960, 5.2) (1990, 5.4) (2010, 9.7)
 (1970, 3.9)



(1900, 47.3) (1930, 59.7) (1960, 69.7) (1990, 75.4)
 (1910, 50.0) (1940, 62.9) (1970, 70.8) (2000, 77.0)
 (1920, 54.1) (1950, 68.2) (1980, 73.1) (2010, 78.7)



(1940, 1315) (1970, 7501) (1990, 25,236)
 (1950, 3180) (1980, 14,610) (2000, 39,237)
 (1960, 4816)



(1950, 30.2) (1980, 30.0) (2000, 35.3)
 (1960, 29.5) (1990, 32.9) (2010, 37.2)
 (1970, 28.1)

Explain 2 Modeling With a Quadratic Function

Example 2 Using the groups of data sets and their scatter plots from Example 1:

Step 1: Choose the data set that appears to be best modeled by a quadratic function.

Explain your choice, whether you think a quadratic model will be a close fit, and whether any other model might possibly be appropriate. What characteristics do you expect the quadratic model will have?

Step 2: Enter the data for your choice into your graphing calculator in two lists, and perform quadratic regression. Then give the model, defining your variables.

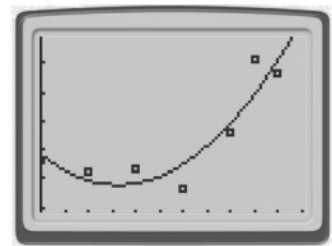
Step 3: Graph the model along with the scatter plot using your calculator, then assess how well the model appears to fit the data.

A Use the data about animal populations in Example 1 Part A.

Step 1: The Florida manatee population appears to be the one best modeled by a quadratic function, as its scatter plot is the only one with a clear change in direction, and it clearly would not be well represented by a linear or exponential model. The whooping crane data might also be fit fairly well on one side of a quadratic model, but it might also be exponential. The quadratic model for the manatee population will have a positive leading coefficient since it opens upward, but it is hard to predict what the y -intercept or the vertex will be. Because the graph is not very symmetrical, the fit may not be very close.

Step 2: A regression model is $y = 48.50x^2 - 317.3x + 3411$ where x is the number of years after 2001 and y is the number of manatees

Step 3: The model is not a close fit, but it does look like an appropriate model for the overall trend during the time of the data. It misses the horizontal position for the vertex by a fairly wide margin, but otherwise is not too far from the data.



- B** Use the data about automobiles in Example 1 Part B.

Step 1:

Step 2:

Step 3:

Reflect

- 5. Discussion** The Florida manatee has been under consideration for being downgraded from endangered to threatened. Do you believe the graph and model of the manatee population in Part A of this Example support this concept or argue against it? Explain your reasoning.

- 6.** How might the model for miles per gallon affect a decision on when to purchase a car?

Your Turn

- Using the data in Your Turn Exercise 4, complete the three steps described at the beginning of Example 2. Also, tell whether you would expect the trend indicated by your model to continue for a time after the data shown, or whether you expect that it would soon change, and explain your answer.

Explain 3 Modeling with an Exponential Function

Example 3 Using the groups of data sets and their scatter plots from Example 1:

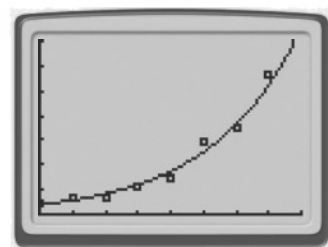
- Step 1:** Choose the data set that appears to be best modeled by an exponential function. Explain your choice, whether you think an exponential model will be a close fit, and whether any other model might possibly be appropriate. What characteristics do you expect the exponential model will have?
- Step 2:** Enter the data for your choice into your graphing calculator in two lists, and perform exponential regression. Then give the model, defining your variables. What are the initial value, growth or decay factor, and growth or decay rate of the model?
- Step 3:** Graph the model along with the scatter plot using your calculator, then assess how well the model appears to fit the data.

- (A) Use the data about animal populations in Example 1 Part A.

Step 1: The whooping crane population appears to be the one best modeled by an exponential function, as it rises increasingly quickly, but does not reflect a change in direction as a quadratic model can. Though the whooping crane plot is nearly linear in its midsection, the slow initial rise and fast later rise indicate that an exponential model is better. The California least tern data show a significant jump, but no clear pattern. An appropriate whooping crane model shows exponential growth, so the parameter b is greater than 1. Because the growth is not very large considering the time period of 70 years, however, the yearly growth factor will not be much above 1.

Step 2: A regression model is $y = 20.05(1.0374)^x$ where x is the number of years after 1940 and y is the population. The initial value for the model is 20 whooping cranes. The growth factor is 1.0374, and the growth rate is 0.0374, or about 3.74% per year.

Step 3: The model is very good fit for the data. Some data points are a little above the curve and some a little below, but the fit is close and reflects the trend very well.



- (B) Use the data about automobiles in Example 1 Part B.

Step 1:

Step 2:

Step 3:

Reflect

- 8. Discussion** What does the model for the whooping crane population predict for the population in 2040? Do you think it is possible that the whooping crane will be removed from the endangered species list any time in the next few decades? Explain.

Your Turn

- 9.** Using the data in Your Turn Exercise 4, complete the three steps described at the beginning of Example 3. Also, tell whether you would expect the trend indicated by your model to continue for a time after the data shown, or whether you expect that it would soon change, and explain your answer.

 **Elaborate**

- 10. Discussion** How does making a prediction from a model help make a decision or judgment based on a given set of data?

11. Describe the process for obtaining a regression model using a graphing calculator.

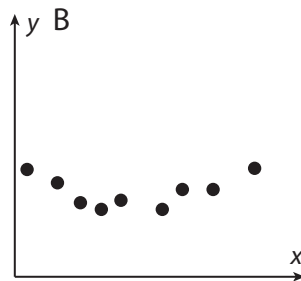
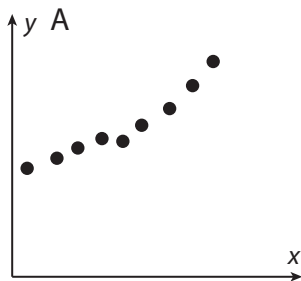
12. **Essential Question Check-In** How can a scatter plot of a data set help you determine the best type of model to choose for the data?

 **Evaluate: Homework and Practice**

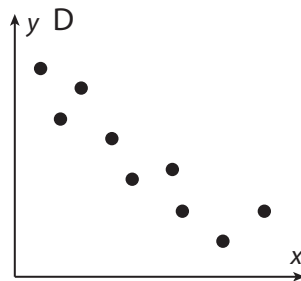
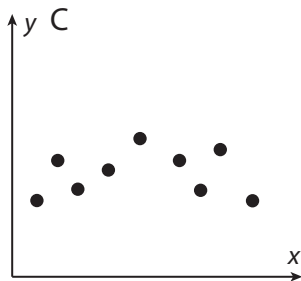


- Online Homework
- Hints and Help
- Extra Practice

1. Match each scatter plot with the most appropriate model from the following. Do not use any choice more than once.



- I. quadratic, $a > 0$
- II. quadratic, $a < 0$
- III. linear, $a < 0$
- IV. exponential, $b > 1$
- V. exponential, $b < 1$



© Houghton Mifflin Harcourt Publishing Company

In Exercises 2–5, for the data set given:

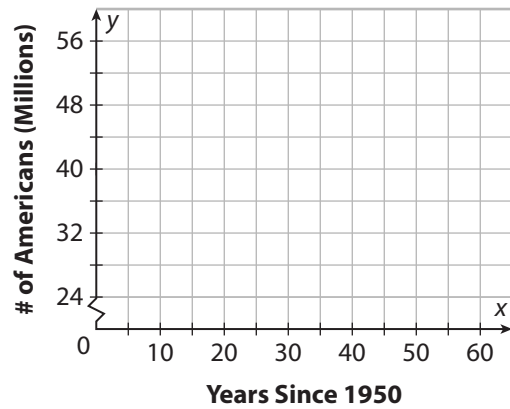
- a. Create a scatter plot of the data. What kind of model do you think is most appropriate? Why? What characteristics do you expect that this model will have?
- b. Use a graphing calculator to find an equation of the regression model. Then interpret the model, including the meaning of important parameters.
- c. Graph the regression model with its scatter plot using a graphing calculator. How well does the model fit?
- d. Answer the question following the data.

2. Population Demographics The data set shows the number of Americans living in multigenerational households.

Year	Number (in Millions)
1950	32
1960	27
1970	26
1980	28
1990	35
2000	42
2010	52



What does the model predict for the number in 2020? in 2040?
Are these numbers reasonable? Explain.



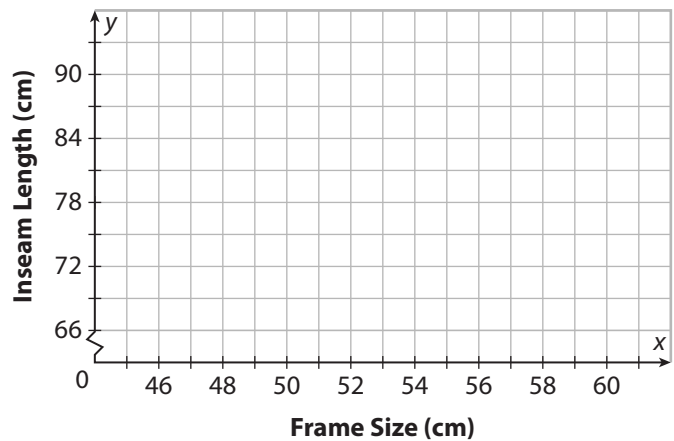
© Houghton Mifflin Harcourt Publishing Company · ©Dean Pictures/Corbis

3. **Cycling** The data set shows the inseam length for different frame sizes for road bicycles.

Frame Size (cm)	Inseam Length (cm)
46	69
48	71
49	74
51	76
53	79
54	81
58	86
60	89
61	91



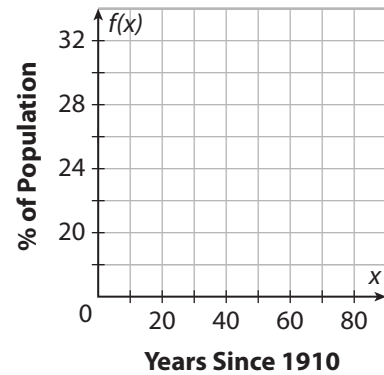
Jarrell has an inseam of 84 cm, but the table does not give a frame size for him. He graphs the model on a graphing calculator and finds that a y -value of 84 is closest to an x -value of 56. He decides he needs a 56 cm frame. Do you think this is a reasonable conclusion. Explain.



- 4. Population Geography** The data set shows the percent of the U.S. population living in central cities.

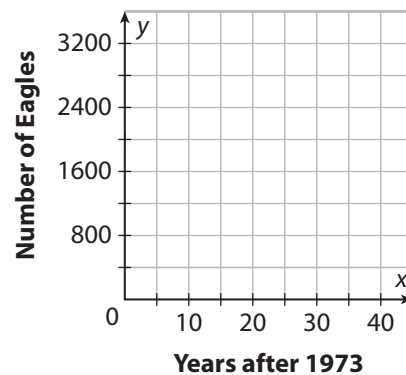
What does your model predict for the percent of the population living in central cities in 2010? How much confidence would you have in this prediction? Explain. Given that the actual number for 2010 was about 36.9%, does this support your judgment?

Year	% of Population
1910	21.2
1920	24.2
1930	30.8
1940	32.5
1950	32.8
1960	32.3
1970	31.4
1980	30.0



- 5. Animal Migration** The data set shows the number of bald eagles counted passing a particular location on a migration route. Predict the number of bald eagles in 2033. How much confidence do you have in this prediction?

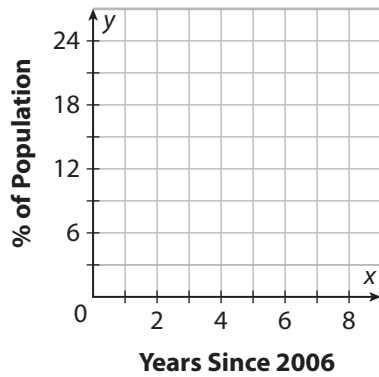
Year	Number of Eagles
1973	41
1978	79
1983	384
1988	261
1993	1725
1998	3289
2003	3356



5. (continued)

6. **Smart Phones** For parts a–d, use the data in the table. The data set shows the percent of the world’s population owning a smart phone.

a. Create a scatter plot of the data.



Year	% of Population
2006	1
2007	3
2008	4
2009	5
2010	7
2011	11
2012	16
2013	22

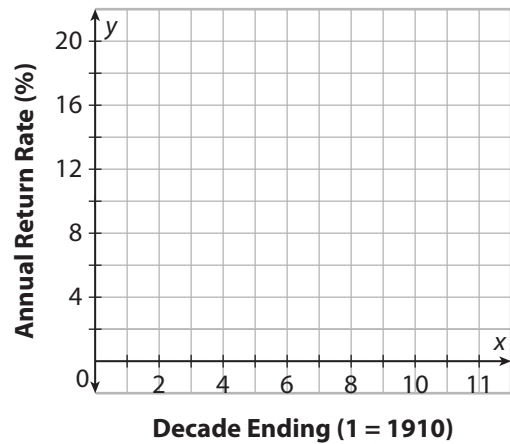
b. Use a graphing calculator to find equations for both exponential and quadratic regression models. Then interpret the models, including the meaning of important parameters.

- c. Graph both regression models with the scatter plot using a graphing calculator. Do both models fit the data well? Does one seem significantly better than the other?

- d. For how long after the data set do you think either model will be a good predictor? Explain your reasoning.

7. Stock Market The data set gives the U.S. stock market's average annual return rates for each of the last 11 decades.

Decade Ending	Annual Return Rate (%)
1910	9.96
1920	4.20
1930	14.95
1940	-0.63
1950	8.72
1960	19.28
1970	7.78
1980	5.82
1990	17.57
2000	18.17
2010	3.1



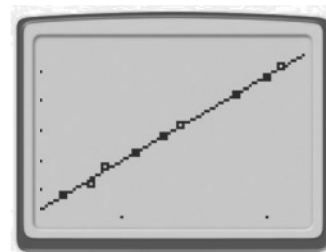
- a. Make a scatter plot of the data.
- b. What kind of model do you think is most appropriate for the data? Explain your reasoning.
- c. Do the data give any kind of prediction about investing in the stock market for the future? Explain.

H.O.T. Focus on Higher Order Thinking

- 8. Explain the Error** Out of curiosity, Julia enters the stock market data from Exercise 7 into her calculator and performs linear and quadratic regression. As she expected, there is almost no fit to the data at all. She then tries exponential regression, and get the message “ERR: DOMAIN.” Why does she get this message?

- 9. Critical Thinking** A student enters the road bicycle data from Exercise 3, accidentally performs quadratic regression instead of linear, and has the calculator graph the regression model with the scatter plot. The student sees this graph:

The model graphed by the calculator is $y = -0.001241x^2 + 1.587x - 1.549$. It is obviously a very close fit to the data, and looks almost identical to the linear model. Explain how this can be true. (*Hint: What happens when you zoom out?*)



- 10. Critical Thinking** A graphing calculator returns a linear regression equation $y = ax + b$, a quadratic regression equation $y = ax^2 + bx + c$, and an exponential regression equation $y = a \cdot b^x$. For exponential regression, a is always positive, which is not true of the other models. How does an exponential model differ from the other two models regarding translations of a parent function?

- 11. Extension** In past work, you have used the correlation coefficient r , which indicates how well a line fits the data. The closer $|r|$ is to 1, the better the fit, and the closer to 0, the worse the fit. When you perform quadratic or exponential regression with a calculator, the *coefficient of determination* r^2 or R^2 serves a similar purpose. Return to the smart phone data in Exercise 6, and perform the quadratic and exponential regression again. (Make sure that “Diagnostics” is turned on in the Catalog menu of your calculator first.) Which model is the closer fit to the data, that is, for which model is R^2 closest to 1?

