4.1-TCU: Investigating Patterns in Data

LEARNING GOALS

Students should understand that:

- Bivariate numerical data may form a nonlinear pattern.
- The value of the parameter *b* in the general form of an exponential model, $\hat{y} = a \cdot b^x$, determines whether the exponential model increases (grows) or decreases (decays).
- The parameter *a* in an exponential model, $\hat{y} = a \cdot b^x$, is the initial value of *y*.
- The parameter *a* in the general form of an exponential equation corresponds to the value of the y-variable when the x-variable is zero.
- The parameter *b* in the general form of an exponential equation provides information about whether the exponential model increases (grows) or decreases (decays).
- The value of *b* provides information about how rapidly the exponential model increases or decreases.
- The value of *b* can be used to identify the percentage change in the *y*-variable for each one-unit increase in the *x*-variable.

Students should be able to:

- Recognize the general form of an exponential model.
- Use an exponential model to make predictions.
- Identify whether an exponential model represents exponential growth or exponential decay.
- Interpret the meaning of a given an exponential model of the form $\hat{y} = a \cdot b^x$.
- Use the value of *b* in an exponential model of the form $\hat{y} = a \cdot b^x$ to determine whether the model represents exponential growth or exponential decay.
- Interpret parameters of an exponential function in context.
- Distinguish between exponential growth and exponential decay using a graph, symbolic form of the function, table of values, or verbal explanation.
- Translate between a percent change and the value of *b* in the exponential model.
- Use an exponential function to make predictions (given *x*, find *y*) graphically and algebraically.
- From an appropriate graph, be able to estimate the initial value parameter. This includes, as necessary, redefining the predictor variable to have *x* = 0 in the data set, or at least not far from the *x*-values in the data.

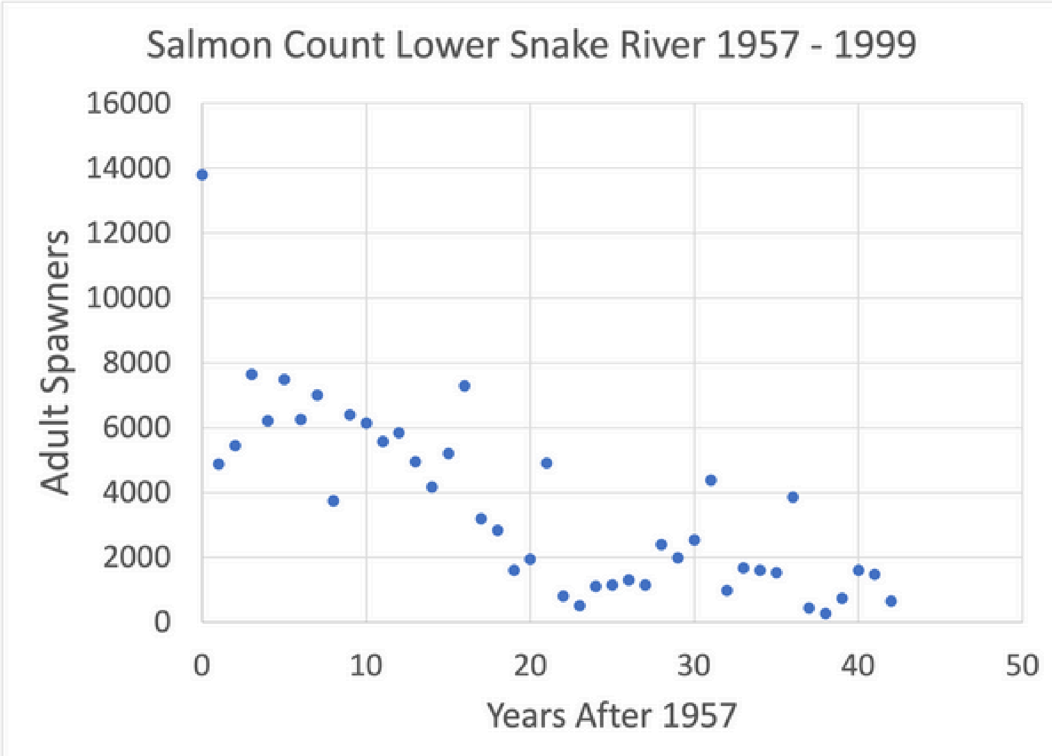
INTRODUCTION

In the year 2000, a report[1] was released describing the decline of salmon populations along the lower Snake River in the northwestern United States. Between 1961 and 1975 four dams were constructed along that river.

Adult salmon swim upriver to the places they were born to spawn. Scientists have methods to obtain approximate counts of the number of adult salmon who do this. You will see a table and a scatterplot that show these counts of adult salmon every year from 1957-1999.

Year	Years after 1957, ***x***	Count
1957	0	13800
1958	1	4870
1959	2	5446
1960	3	7634
1961	4	6212
1962	5	7478
1963	6	6249
1964	7	7003
1965	8	3735
1966	9	6391
1967	10	6134
1968	11	5571
1969	12	5842
1970	13	4952
1971	14	4171
1972	15	5203
1973	16	7286
1974	17	3183
1975	18	2835
1976	19	1605
1977	20	1936
1978	21	4904
1979	22	807
1980	23	513
1981	24	1101
1982	25	1150
1983	26	1307
1984	27	1150
1985	28	2395
1986	29	1978
1987	30	2538
1988	31	4385
1989	32	978
1990	33	1672
1991	34	1595
1992	35	1532
1993	36	3857
1994	37	438
1995	38	262
1996	39	733
1997	40	1592
1998	41	1472
1999	42	652

[1] http://www.life.illinois.edu/ib/463/IB463files/paperstoread/kareiva.pdf



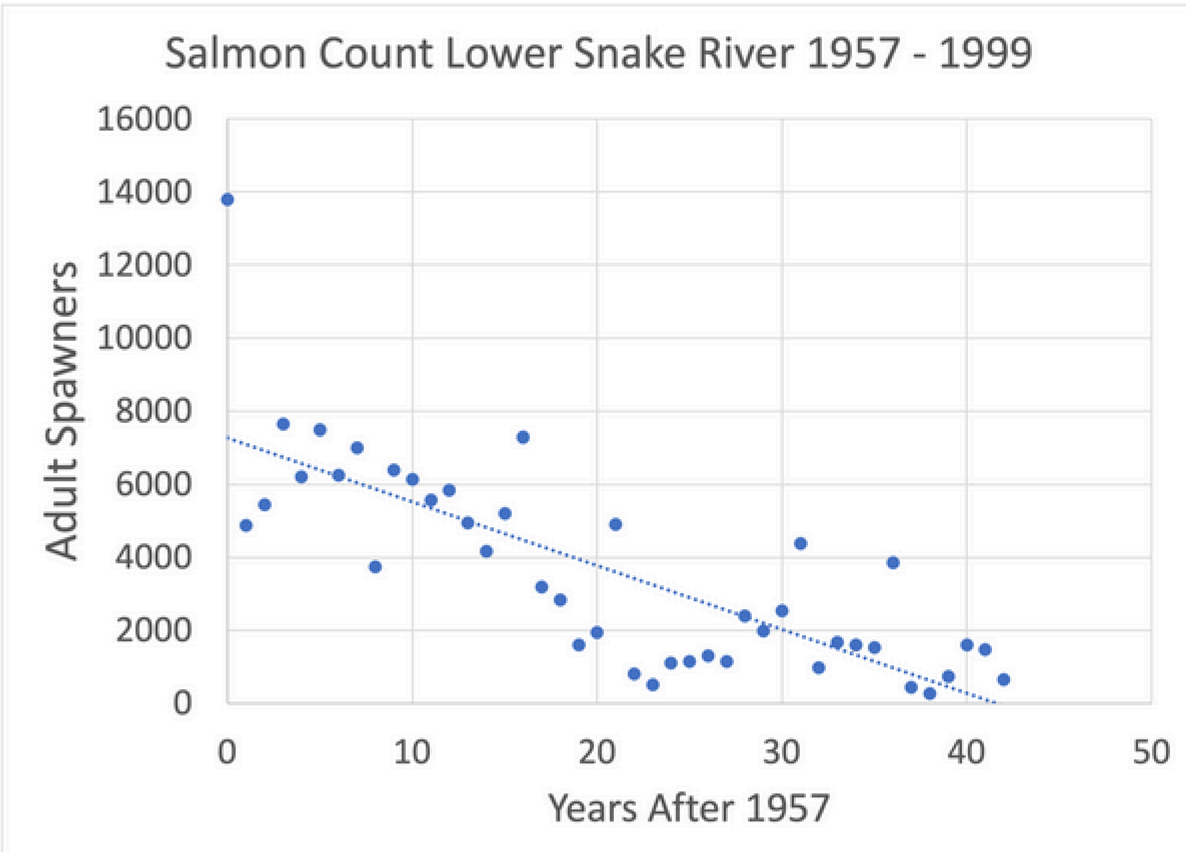
1 Use the raw data and/or scatterplot to answer the following questions:

A The mean adult count for 1957-1999 is 3594 adult salmon. In what year did the adult salmon count become consistently below this value?

B In what year did the adult salmon count first dip below 1000?

TRY THESE

We now use a least-squares line to summarize the adult salmon count for the years after 1957. Below are the scatterplot and corresponding least-squares regression line.



Least-squares Regression Line: $\hat{y} = -174.8x + 7265$

2 Use the information above to answer the following questions:

A Look at the least-squares equation. What is the meaning of the number 7265?

B What is the meaning of the number −174.8 in the least-squares equation?

C According to the regression line, when did the salmon count reach 3594? Round to the nearest whole number.

D According to the regression line, when did the salmon count reach 2835? Round to the nearest whole number.

E According to the regression line, when did the salmon count reach 0? *Round to the nearest whole number.*

3 Do you think the regression line is a good summary of the data? Why or why not?

NEXT STEPS

Nonlinear Models

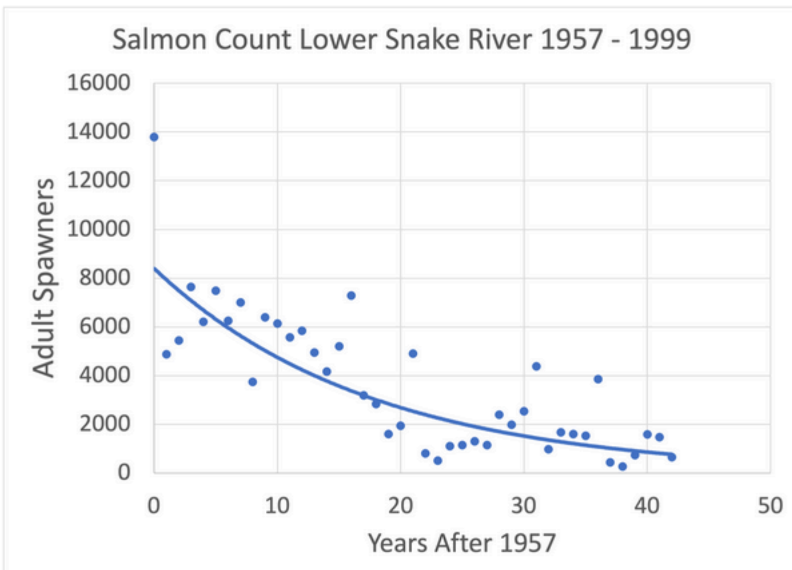
There are two main types of least-squares regression models: linear models and non-linear models. Earlier in Module 6, we studied a linear model and *least-squares regression lines*. Linear models are often useful, but they don't represent all real world behavior.

Sometimes we use least-squares methods to develop *non-linear mathematical models* that summarize relationships between variables.

One type of non-linear model is called an **exponential** model. An exponential model has the form: $\widehat{y} = a \cdot b^{x}$. We use exponents to indicate repeated multiplication. Later in this Collaboration, we will discuss exponential models in more detail and will provide an example.

Language Tip: An **exponential model** is represented by an equation with a variable as an exponent.

In the scatterplot on the next page, we fit an exponential model to the data. This model describes the relationship between the variables very well. There is very little scatter about the exponential model, so there is a strong **exponential relationship** between these variables. Below the graph is the equation of the exponential model.



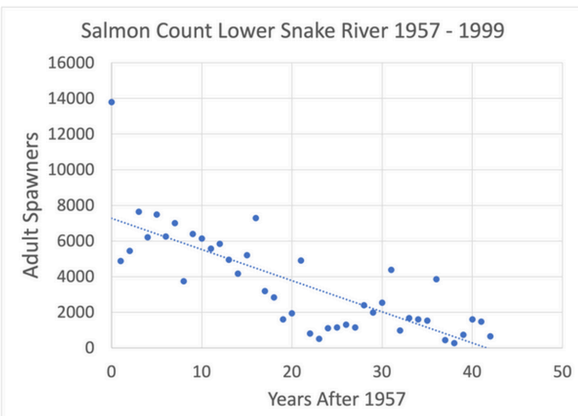
Exponential Model: $\widehat{y} = 8396(0.94)^{x}$

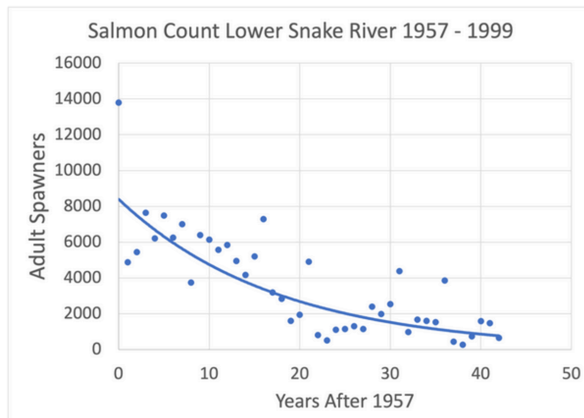
4 In previous units, you learned that the *coefficient of determination*, r^2, is useful because it allows us to assess fit for linear *and* non-linear models. Remember that in a least-squares model r^2 is the proportion of total variability in the *y*-variable that can be explained by the *x*-variable.

A For the linear model, $r^2 \approx 0.616$. Interpret this value using the information you have about the adult salmon count.

B For the exponential model, $r^2 \approx 0.605$. Interpret this value using the information you have about the adult salmon count.

5 While the coefficient of determination is a little higher for the linear model, r^2 is a similar value for both the linear and exponential models. Another factor to consider is how your model will perform in predicting future values. In this case, predicting adult salmon counts in years after 1999, which is the last year we have data.





$$\widehat{y} = -174.8x + 7265$$

$$\widehat{y} = 8396(0.94)^x$$

A Use the **linear model** to predict the salmon count for the year 2010. Round your answer to the nearest whole salmon.

B Use the **exponential model** to predict the salmon count for the year 2010. Round your answer to the nearest whole salmon.

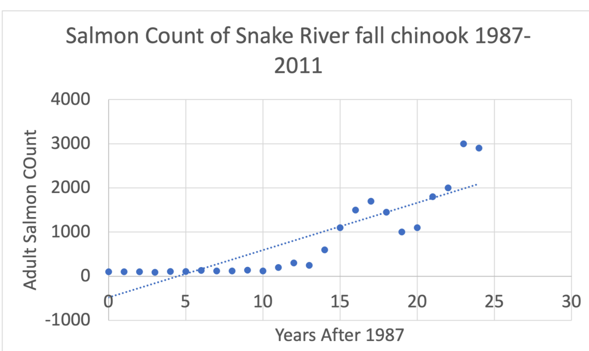
C Which do you think is more accurate? Please explain your reasoning.

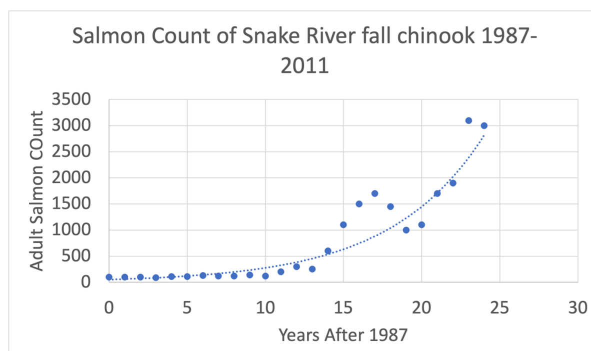
MAKING A COMEBACK

Fortunately, we never had to see the adult salmon count reach zero. There have been many success stories in rebuilding the salmon population. Interventions have included creating paths for salmon to get around dams, dam removal and creating salmon hatcheries along these rivers.

One particularly successful intervention was led by the Nez Perce Tribe in collaboration with The National Oceanic and Atmospheric Administration (NOAA). In 1992, the NOAA listed Snake River fall chinook salmon as an endangered species. This designation threatened the tribal fishery of the Nez Perce. The tribe worked with multiple organizations to develop “a cutting-edge hatchery program that allows the Nez Perce Tribe to supplement natural chinook populations with hatchery-reared fish of the same stock.”[2]

Data from the years 1987-2011 show a substantial increase in the adult count of Snake River fall chinook. Data is shown below with both linear and exponential models.





$$\widehat{y} = 109x - 494.6$$

$$\widehat{y} = 41(1.20)^{x}$$

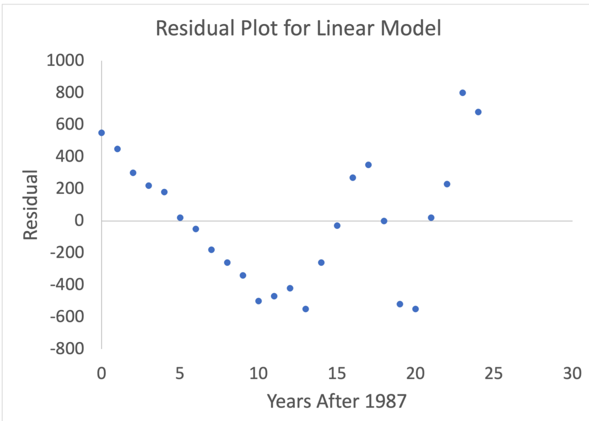
6 For the linear model, $r^2 \approx 0.778$, and for the exponential model, $r^2 \approx 0.873$. Interpret what this says about the effectiveness of the two models.

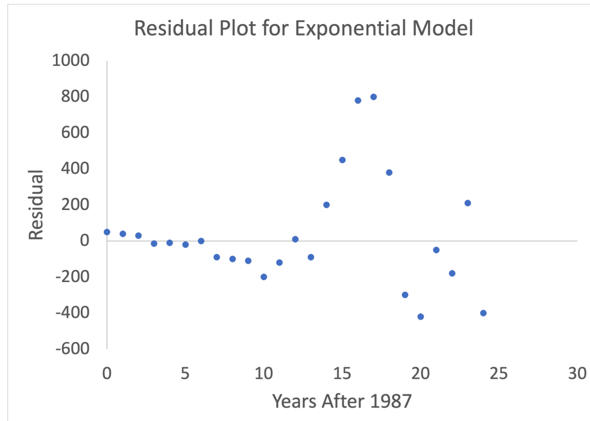
[2] https://critfc.org/wp-content/uploads/2012/11/success-snake_fall_chinook.pdf

Another way to study the effectiveness of a model is to look at residual plots. Recall that residuals are calculated by subtracting the model's prediction from the actual data.

$$residual = y - \hat{y}$$

Below are residual plots for both the linear and exponential models:





7 Compare the residual plots above. Do the residual plots indicate that a linear model or an exponential model is an appropriate fit for the data? Why do you think so?

8 In the both the linear and exponential models, x represents the number of years since 1987 and $\hat{y}$ represents the predicted count for adult salmon in the Snake River.

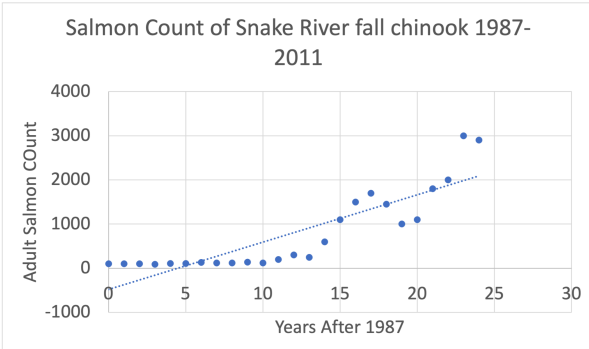
A Complete the left column in the table by filling in the appropriate x value for each year:

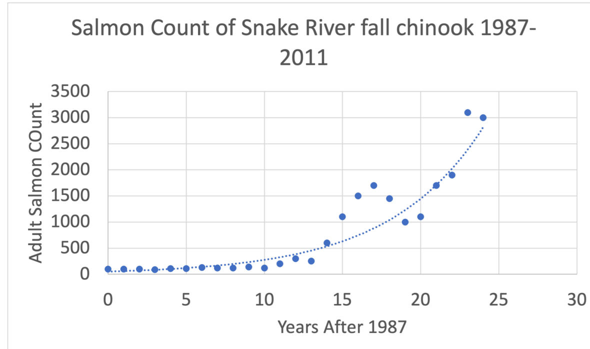
B Find predictions for each year using the linear model.

C Find predictions for each year using the linear model. Round your answers to one decimal place.

Year	**Years since 1987, x**	**Prediction using linear model** $\hat{y} = 109x - 494.6$	**Prediction using exponential model** $\hat{y} = 41(1.20)^{x}$
1987			
1988			
1995			
2000			
2010			
2020			

Here, again is the data from 1987-2011





$\widehat{y} = 109x - 494.6$

$\widehat{y} = 41(1.20)^x$

9 Use the models to visually estimate when the adult salmon count would surpass 4000. Which do you expect to be more accurate?

BREAKING DOWN THE MODEL [20 MIN]

Earlier in this Collaboration, we explained that an exponential model is one type of non-linear model. An exponential model has the form: $\widehat{y} = a \cdot b^x$. We use exponents to indicate repeated multiplication. For example, $8^3 = 8 \cdot 8 \cdot 8 = 512$. Thus b is repeatedly multiplied in an exponential model. It does not make any sense to multiply repeatedly by negative numbers. It also does not make any sense to multiply by 0 or by 1. For this reason b must be positive and not equal to 1.

When we multiply a positive number by a number between 0 and 1, like 1/2, the answer is always smaller than the original number. For example, 10 · (1/2) = 5. When we multiply a positive number by a number greater than 1, like 2, the answer is always greater than the original number. For example, 10 · 2 = 20. So when b is between 0 and 1, we say the model represents **exponential decay**. When b is greater than 1, we say the model represents **exponential growth**.

Recall that when we use 0 as an exponent for any positive number, the result is 1. For example, $8^0 = 1$. So when $x = 0$, $\widehat{y} = a \cdot b^0 = a \cdot 1 = a$. We call a the initial value of the exponential model. It is also the y-intercept of the graph of the model.

YOU NEED TO KNOW

An exponential model has the general form: $\widehat{y} = a \cdot b^x$

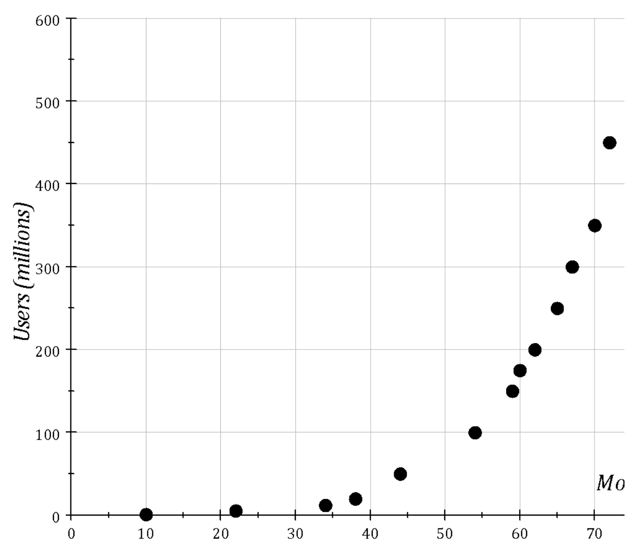
The numbers a and b in the exponential model have the following key properties:

- a is the y-intercept of the model. It is the y-value when $x = 0$. It is also the initial value.
- $b > 0$ and $b \neq 1$.
- If $b > 1$, b is the **growth factor**, and the model represents **exponential growth**.
- If $b < 1$, b is the **decay factor**, and the model represents **exponential decay**.

Facebook Membership

The online social networking site Facebook was launched in February of 2004. From 2004 to 2010, Facebook membership grew rapidly as its popularity increased. Facebook tracks its number of users closely and uses this data to generate new business opportunities. The table and graph below provide information about Facebook membership (measured by number of Facebook users) for selected months between December 2004 and July 2010. The membership counts represent the number of people who are Facebook users. The numbers are given in millions, so 5.5 means 5,500,000.[3]

Month-Year	Months since February 2004, *x*	Facebook Actual Membership (millions of users)
Dec-04	10	1
Dec-05	22	5.5
Dec-06	34	12
Apr-07	38	20
Oct-07	44	50
Aug-08	54	100
Jan-09	59	150
Feb-09	60	175
Apr-09	62	200
Jul-09	65	250
Sep-09	67	300
Dec-09	70	350
Feb-10	72	450
Jul-10	77	500



[3] http://www.benphoster.com/facebook-user-growth-chart-2004-2010/

Facebook membership can be modeled by the exponential curve

$$\hat{y} = 0.8 \cdot 1.09^{x}$$

In this equation, $\hat{y}$ is the predicted number of Facebook users (in millions) and x is the number of months since Facebook was launched in February of 2004.

10 According to the exponential model, when did Facebook's *predicted* membership exceed (grow larger than) 200 million users?

11 What do the values 0.8 and 1.09 in the exponential model mean? In your explanation, think about the exponential curve.

12 Use the exponential model to predict Facebook's membership in September 2007 (month 43). Round to the nearest million users.

13 Use the exponential model to predict Facebook's membership in January 2010 (month 71). Round to the nearest million users.

14 Use the exponential model to predict Facebook membership in March 2011 (month 85). Round to the nearest million users.

15 Consider the predictions you just made in your answers to Questions 10 and 11. Which do you think is more reliable? Give a reason for your answer. Think about the data that were used to create the model.

16 Membership data was not reported until eight months after Facebook's launch.

A Use the exponential model to predict how many users Facebook had when it started in February of 2004 (month 0)? Do not round your answer.

B Do you think this answer is realistic? Why or why not?

We now focus on the meaning of *b* in exponential growth and exponential decay models. In the Facebook membership exponential model, *b* equals 1.09:

$$\hat{y} = 0.8 \cdot 1.09^x$$

Since $b > 1$, the equation models exponential growth.

17 Use the Facebook membership exponential model to answer the following questions.

A What does the model predict for Facebook membership in May 2007? Round to two decimal places.

B What does the model predict for Facebook membership in June 2007? Round to two decimal places.

C Divide the predicted membership in June 2007 by the predicted membership in May 2007. Round your answer to two places after the decimal.

$$\frac{Predicted\ June\ 2007\ membership}{Predicted\ May\ 2007\ membership} =$$

D What do you notice about the answer to the last question? (*Hint*: Think about the exponential growth model.)

E What was the percent increase in Facebook membership between May 2007 and June 2007? In your answer, round each value to two decimal places. Remember that percent increase and percent decrease are always the change divided by the original amount. Round to the nearest percent.

$$\%\ increase = \frac{(June\ 2007\ membership) - (May\ 2007\ membership)}{(May\ 2007\ membership)} \cdot 100\% =$$

In this example, the growth rate is 9%. This means that the model predicts that Facebook membership increases by 9% each month. This results in a *growth factor* of 1.09. The percent growth, when written as a decimal, is equal to the growth factor minus 1, or, in other words, 1.09 – 1 = 0.09.

YOU NEED TO KNOW

An exponential model indicates that, for every one unit increase in *x*, the value of $\hat{y}$ increases or decreases by a fixed percent. The value of b, the growth or decay factor, is related to the percent increase or decrease in the following way:

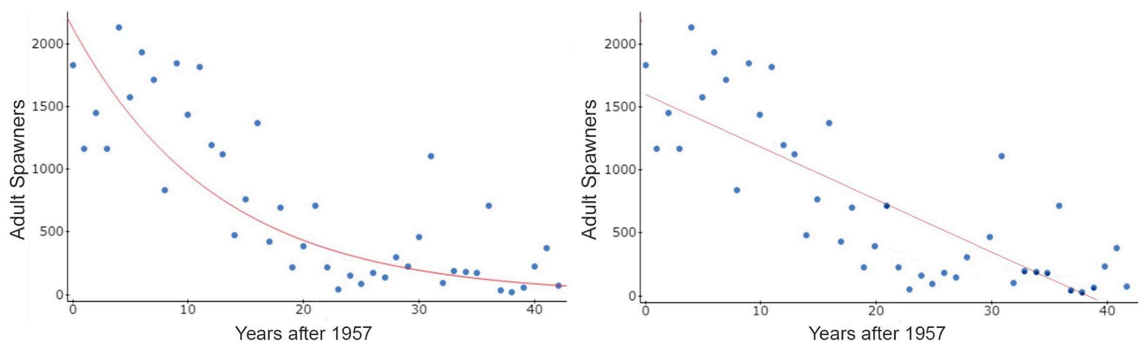
$b - 1$ = growth or decay rate as a decimal

Multiply by 100% to change the decimal rate into a percent.

Exercise 6.7-TCU

The scatterplots below show the adult salmon counts for a subsection of the Snake River, called Bear Creek, from 1957 to 1999. One scatterplot is fitted with a linear model and the other is fitted with an exponential model. The equations of the models are given below each scatterplot. In both equations, x represents the number of years since 1957. Use the graphs and equations to answer the following questions.

Adults Spawners in Bear Creek



Exponential Model: $\widehat{y} = 2077(0.925)^x$ Linear Model: $\widehat{y} = -41.6x + 1602$

Complete the missing entries, Questions 1-3, in the table below:

Year	Years since 1957, x	Predicted Adult Salmon Count (Linear)	Predicted Adult Salmon Count (Exponential)
1960	3	1477.2	1643.8
1970	13	1061.2	753.8
1980	(1)	(2)	(3)
1990	33	229.2	158.5

4 Which statement best explains the meaning of 1602 in the linear model?

(i) Each year, the salmon population is predicted to increase by 1602.
(ii) Each year, the salmon population is predicted to decrease by 1602.
(iii) The predicted salmon count in 1957 was 1602.
(iv) The model predicts it will take 1602 years for the salmon count to reach 0.

5 Which statement best explains the meaning of -41.6 in the linear model?

(i) Each year, the salmon population is predicted to increase by 41.6.
(ii) Each year, the salmon population is predicted to decrease by 41.6.
(iii) The predicted salmon count in 1957 was 41.6.
(iv) The model predicts it will take 41.6 years for the salmon count to reach 0.

6 Explain the meaning of 2077 in the exponential model. In your explanation, think about the predicted radiation level.

(i) Each year, the salmon population is predicted to increase by 2077.
(ii) Each year, the salmon population is predicted to decrease by 2077.
(iii) The predicted salmon count in 1957 was 2077.
(iv) The model predicts it will take 2077 years for the salmon count to reach 0.

7 Explain the meaning of 0.925 in the exponential model. In your explanation, think about the predicted radiation level.

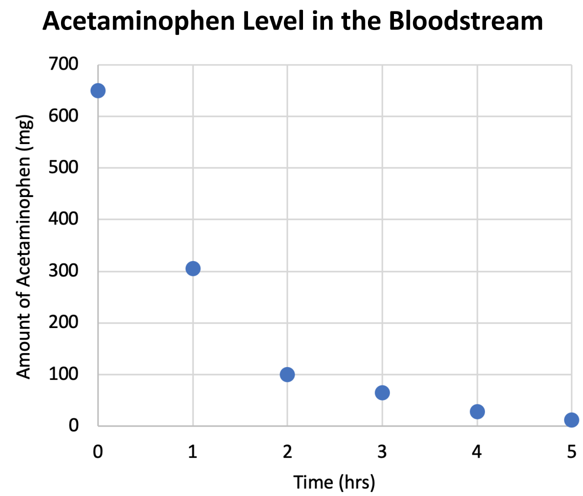
(i) Each year, the salmon count is predicted to increase by 92.5%.
(ii) Each year, the salmon count is predicted to increase by 7.5%.
(iii) Each year, the salmon count is predicted to decrease by 92.5%.
(iv) Each year, the salmon count is predicted to decrease by 7.5%.

8 When does the **linear** model predict the adult salmon count will reach 100? Round to the nearest year.

9 When does the **exponentia**l model predict the adult salmon count will reach zero? Round to the nearest year.

Most oral medications take effect by entering our bloodstream. During the hours after we swallow the initial dose of the medicine, the amount of it in our bloodstream slowly decreases. We can model this decrease with an exponential curve. The following table shows the amount (in mg) of acetaminophen (Tylenol is one common brand name for acetaminophen) in a person's bloodstream h hours after a person takes the medicine. Use the table and graph below to answer the following questions.

Hours after Swallowing, h	Acetaminophen in the Bloodstream (in mg)
0	650
1	305
2	100
3	65
4	28
5	12



10 If the exponential model $\hat{y} = 650 \cdot 0.46^h$ predicts the amount of acetaminophen (in mg) in a person's bloodstream h hours after taking 650 mg of the drug, what is the initial amount of acetaminophen in the person's bloodstream?

11 When will the amount of acetaminophen be below 90 mg? Write whole numbers below.

Sometime between __________ and __________ hours after taking the acetaminophen.

12 The exponential curve $\hat{y} = 650 \cdot 0.46^h$ predicts the amount of acetaminophen in a person's bloodstream h hours after taking the drug. Use this equation to predict the amount of acetaminophen in a person's bloodstream 3.5 hours after taking the medicine. Round your answer to one place after the decimal.

13 What is the value of b in the exponential equation?

14 Does this value of b indicate exponential growth or exponential decay?

(i) 0.46 is positive. This indicates growth.
(ii) 0.46 is more than 1. This indicates growth.
(iii) 0.46 is less than 1. This indicates 46% decay.
(iv) 0.46 is less than 1. This indicates 54% decay.

15 Determine the percent decrease in the amount of acetaminophen in a person's blood during the third hour after taking the medicine. Round to the nearest percent.

$$\%\ decrease = \frac{(amount\ at\ h = 3) - (amount\ at\ h = 2)}{(amount\ at\ h = 2)} \times 100 =$$

The number of people in the United States who are 65 years and older is projected to increase rapidly over the next few decades. In 2010, there were approximately 40 million people age 65 years and older in the U.S. (Source: 2010 U.S. Census). Statisticians have projected that this number will increase to 89 million by 2050.[4] Assuming that population growth is exponential, the U.S. population that is 65 years and older (in millions) can be modeled by

$$\hat{y} = 40 \cdot 1.021^{x}$$

where x is the number of years after 2010.

16 What is the predicted number of people in the U.S. age 65 years and older in 2012? Round your answer to two decimal places.

17 What is the predicted number of people in the U.S. age 65 years and older in 2030? Round your answer to two decimal places.

18 What is the percent increase in the number of people in the U.S. age 65 and older each year?

[4] http://www.census.gov/prod/2010pubs/p25-1138.pdf