Exercises

3.2.1 For each of the data sets below, determine which model - exponential, power, or logarithmic - "best" fits the data.

a.

x	1	2	3	4	5	6
y	1.66	2.41	6.04	9.89	17.31	31.54

b.

x	1	2	3	4	5	6
y	2.68	5.61	8.71	9.83	11.16	11.03

3.2.2 Table 3.2 contains the total length and weight of 20 black bears (data collected by Brett Troyer, 2011). Graph weight vs length, fit different linearizable models to the data, and select the one that "best" fits the data.

Weight	110	60	90	60	85	95	85	155	220	105
Length	139	138	139	120.5	149	141	150	166	180	129.5
Weight	110	115	255	105	90	75	75	115	35	140
Length	150	142	162	148	140	134	137	149	102	151.5

Table 3.2

3.2.3 For each set of data below, fit a model of the given form by transforming the data appropriately and fitting a straight line to the transformed data. Graph the residuals and analyze how well the model fits the data.

a. Model: $y = ax^2 + b$

x	1	2	3	4	5	6
y	16.3	23.1	37.4	46.9	58.7	91.0

b. Model: $y = a\sin(x) + b$

x	1	2	3	4	5	6
y	1.34	1.61	-0.98	-3.80	-4.55	-2.30

c. Model: $y = a\dfrac{x^2+1}{\ln(x)} + b$

x	2	3	4	5	6	7
y	3.30	5.63	9.52	14.31	19.84	26.061

3.2.4 Consider the data below:

x	0	2	4	6	8	10
y	3.000	3.061	3.122	3.186	3.250	3.316

a. Fit a linear model to the data (a model of the form $y = mx + b$). How well does the model appear to fit the data? Create a graph of the residuals. What do you notice about the pattern of the residuals?

b. Fit an exponential model to the data. How well does the model appear to fit the data? Calculate and graph the residuals. What does this tell you about how well this model fits the data?

3.2.5 Use algebra or the natural logarithm to derive the given "linearization" of each of the following models:

	Model	**Linearization**
a.	$y = \dfrac{1}{a+bx}$	$\dfrac{1}{y} = a + bx$
b.	$y = \dfrac{x}{a+bx}$	$\dfrac{1}{y} = \dfrac{a}{x} + b$
c.	$y = 1 + ae^{bx}$	$\ln(y-1) = \ln a + bx$
d.	$y = 1 - e^{-x^a/b}$	$\ln\left[\ln\left(\dfrac{1}{1-y}\right)\right] = a\ln x + \ln\dfrac{1}{b}$

3.2.6 Growing populations typically have a maximum size called the *carrying capacity*. The size of a population y at time t can often be modeled with the *logistic model*

$$y = \frac{L}{1 + e^{a+bt}}$$

where L is the carrying capacity and a and b are parameters.

a. "Lizearize" this model with simple algebra and the natural logarithm by showing that

$$\ln\left(\frac{L-y}{y}\right) = a + bt.$$

b. The data below show the size of a bacteria population at certain points in time. Assuming the carrying capacity is $L = 625$, fit a logistic model to the data using the linearization in part a.

t	0	1	2	3	4	5	6	7	8	9
y	10.3	17.2	27	45.3	80.2	125.3	176.2	255.6	330.8	390.4
t	10	11	12	13	14	15	16	17	18	19
y	440	520.4	560.4	600.5	610.8	614.5	618.3	619.5	620	621

c. Create a graph of the residuals for the model found in part b. What does this say about the quality of the fit of this model?

Directions for exercises 3.2.7 - 3.2.9: In this section, we found the values of the parameters in the models by transforming the data appropriately and then fitting a straight line to the transformed data. Another approach is to find formulas for the parameters using the least-squares criterion. That is, if the model is $y = f(x)$, then the parameters should minimize the quantity

$$S = \sum_{i=1}^{n} (y_i - f(x_i))^2.$$

This can be done, in principle, by taking the partial derivatives of S with respect to each of the parameters, setting the derivatives equal to 0, and solving for the parameters, much like we did in section 2.4.

In exercises 3.2.7 - 3.2.9 you are given a type of linearizable model and formulas for the parameters. For each exercise:

a. Design a spreadsheet to show, by example, that the given formulas give the same values of the parameters as we would get if we transformed the data appropriately and fit a straight line to the transformed data.

b. Derive the given formulas using the approach described above.

3.2.7 Model: $y = ax^3$

$$a = \frac{\sum y_i x_i^3}{\sum x_i^6}$$

3.2.8 Model: $y = ax^2 + b$

$$a = \frac{-n \sum x_i^2 y_i + \sum y_i \sum x_i}{\left(\sum x_i^2\right)^2 - n \sum x_i^4}, \quad b = \frac{\sum y_i - a \sum x_i^2}{n}$$

3.2.9 Model: $y = a + b \ln x$

$$b = \frac{\sum y_i \sum \ln x_i - n \sum y \ln x_i}{\left(\sum \ln x_i\right)^2 - n \sum \left(\ln x_i\right)^2}, \quad a = \frac{\sum y_i - b \sum \ln x_i}{n}$$

3.3 Coefficient of Determination

The coefficient of determination, denoted R^2, is a numerical measure of how well a line fits a set of data. To illustrate the fundamental concepts, we will generate some hypothetical data according to the relationship $y = 3 + 2x$.