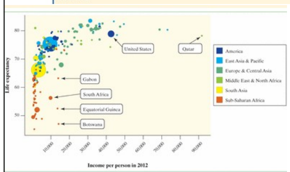
Ch 12.2 Transforming Linearity

Applying a function such as the logarithm or square root to a quantitative variable is called **transforming** the data. We will see in this section that understanding how simple functions work helps us choose and use transformations to straighten nonlinear patterns.

Example 5 Health and Wealth

Straightening out a curved pattern

The Gapminder Web site, www.gapminder.org, provides loads of data on the health and well-being of the world's inhabitants. Figure 12.9 on the next page is a scatterplot of data from Gapminder.[14] The individuals are all the world's nations for which data are available. The explanatory variable is a measure of how rich a country is: income per person. The response variable is life expectancy at birth. We expect people in richer countries to live longer because they have better access to medical care and typically lead healthier lives. The overall pattern of the scatterplot does show this, but the relationship is not linear. Life expectancy rises very quickly as income per person increases and then levels off. People in very rich countries such as the United States live no longer than people in poorer but not extremely poor nations. In some less wealthy countries, people live longer than in the United States.



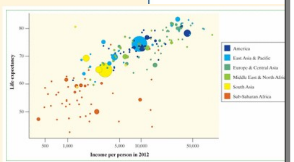


Figure 12.10 Scatterplot of life expectancy against income per perso (on a logarithm scale) for many nations.

Mar 7-12:38 PM

Example 6 Go Fish!

Transforming with powers

Imagine that you have been put in charge of organizing a fishing tournament in which prizes will be given for the heaviest Atlantic Ocean rockfish caught. You know that many of the fish caught during the tournament will be measured and released. You are also aware that using delicate scales to try to weigh a fish that is flopping around in a moving boat will probably not yield very accurate results. It would be much easier to measure the length of the fish while on the boat. What you need is a way to convert the length of the fish to its weight.

You contact the nearby marine research laboratory, and they provide reference data on the length (in centimeters) and weight (in grams) for Atlantic Ocean rockfish of several sizes.[15]

Length:	5.2	8.5	11.5	14.3	16.8	19.2	21.3	23.3	25.0	26.7
Weight:	2	8	21	38	69	117	148	190	264	293
Length:	28.2	29.6	30.8	32.0	33.0	34.0	34.9	36.4	37.1	37.7
Weight:	318	371	455	504	518	537	651	719	726	810

Figure 12.11 is a scatterplot of the data. Note the clear curved shape.

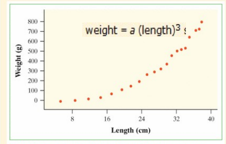


Figure 12.11 Scatterplot of Atlantic Ocean rockfish weight versus length.

This transformation of the explanatory variable helps us produce a graph that is quite linear.

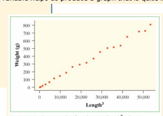


Figure 12.12 The scatterplot of weight versus length^3 is linear.

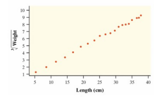


Figure 12.13 The scatterplot of $\sqrt[3]{\text{weight}}$ versus length is linear.

Mar 7-12:45 PM

Example 7 Go Fish!

Transforming with Powers and Roots

Here is Minitab output from separate regression analyses of the two sets of transformed Atlantic Ocean rockfish data.

Transformation 1: (length^3, weight)

Predictor	Coef	SE Coef	T	P
Constant	4.066	6.902	0.59	0.5
Length^3	0.0146774	0.0002404	61.07	0.000

S = 18.8412 R-Sq = 99.5% R-Sq(adj) = 99.5%

Transformation 2: (length, $\sqrt[3]{\text{weight}}$)

Predictor	Coef	SE Coef	T	P
Constant	−0.02204	0.07762	−0.28	0.780
Length	0.246616	0.002868	86.00	0.000

S = 0.124161 R-Sq = 99.8% R-Sq(adj) = 99.7%

(a) Give the equation of the least-squares regression line. Define any variables you use.

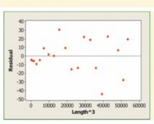
(b) Suppose a contestant in the fishing tournament catches an Atlantic Ocean rockfish that's 36 centimeters long. Use the model from part (a) to predict the fish's weight. Show your work.

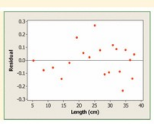
(a) Transformation 1: $\widehat{\text{weight}} = 4.066 + 0.0146774\,(\text{length}^3)$

(b) Transformation 2: $\widehat{\text{weight}} = 4.066 + 0.0146774(36^3) = 688.9$ grams

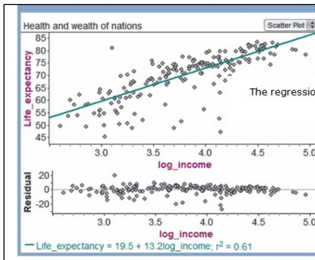
Transformation 2: $\widehat{\sqrt[3]{\text{weight}}} = -0.02204 + 0.246616\,(\text{length})$

$\widehat{\text{weight}} = 8.856^3 = 694.6$ grams





Mar 7-12:48 PM



The regression line is

predicted life expectancy = 19.5 + 13.2 log(income)

predicted life expectancy = 19.5 + 13.2 log(42,296.20) = 80.567 years

income per person of $42,296.20,

Mar 7-12:51 PM

Power Model vs Exponential Model

Power model

(log x, log y) or

(log explanatory variable, response variable)

If the scatterplot of logarithm (or natural logarithm) of the response variable values and the logarithm (or natural logarithm) of the explanatory values has a linear form, then the 2 variables can be modeled using a power function.

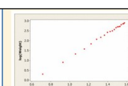
Exponential Model

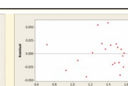
(x, log y) or

(explanatory x, log response variable)

If the scatterplot of logarithm (or natural logarithm) of the response variable values and the original explanatory values has a linear form, then the 2 variables can be modeled using an exponential function.

Mar 7-12:55 PM





Regression Analysis: log(Weight) versus log(Length)

Predictor	Coef	SE Coef	T	P
Constant	−1.89940	0.03799	−49.99	0.000
log(Length)	3.04942	0.02764	110.31	0.000

S = 0.0281823 R-Sq = 99.9% R-Sq(adj) = 99.8%

(a) Based on the output, explain why it would be reasonable to use a power model to describe the relationship between weight and length for Atlantic Ocean rockfish. Power Model - taking logs of both variables.

(b) Give the equation of the least-squares regression line. Be sure to define any variables you use.

(b) $\widehat{\log(\text{weight})} = -1.89940 + 3.04942\log(\text{length})$

Example 9 Go Fish!

Making predictions

PROBLEM: Suppose a contestant in the fishing tournament catches an Atlantic Ocean rockfish that's 36 centimeters long. Use the model from part (b) of the previous example to predict the fish's weight. Show your work.

SOLUTION: For a length of 36 centimeters, we have

$\widehat{\log(\text{weight})} = -1.89940 + 3.04942\log(36) = 2.8464$

To find the predicted weight, we use the definition of a logarithm as an exponent:

$\log_{10}\widehat{(\text{weight})} = 2.8464$

$\widehat{\text{weight}} = 10^{2.8464} \approx 702.1$

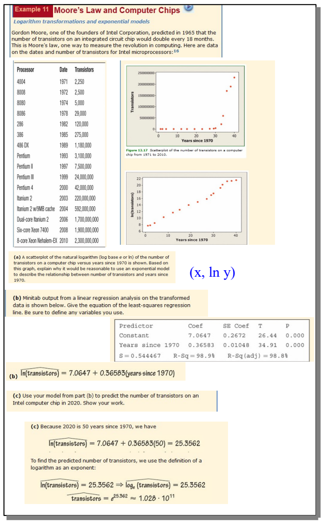
This model predicts that a 36-centimeter-long rockfish will weigh about 702 grams.

$\log_{\text{base}}$ Answer = exponent

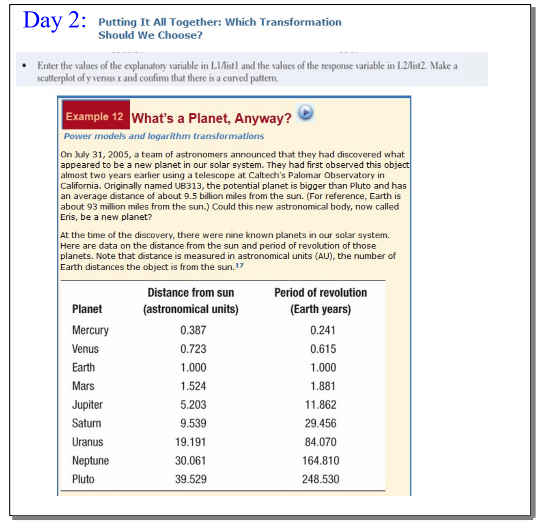
log base 10

ln is base e

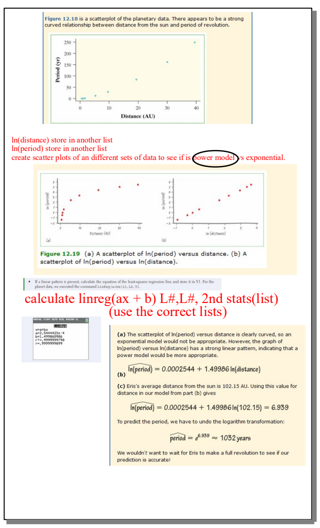
Mar 7-12:53 PM



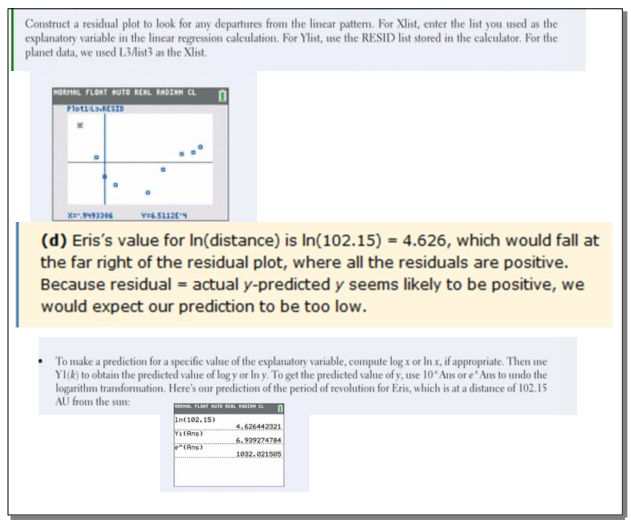
Mar 7-1:08 PM



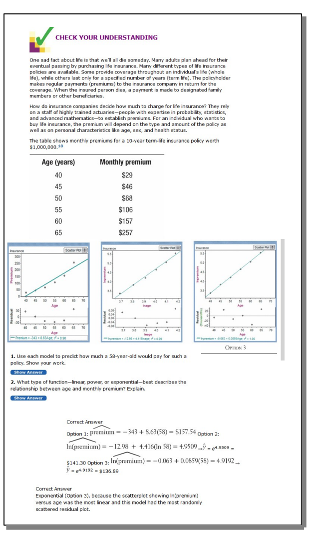
Mar 8-12:07 PM



Mar 8-10:27 AM



Mar 8-12:12 PM



Mar 8-3:25 PM