Making a Linear Model from Data

A <u>math model</u> uses data to construct equations that represent a real life situation.
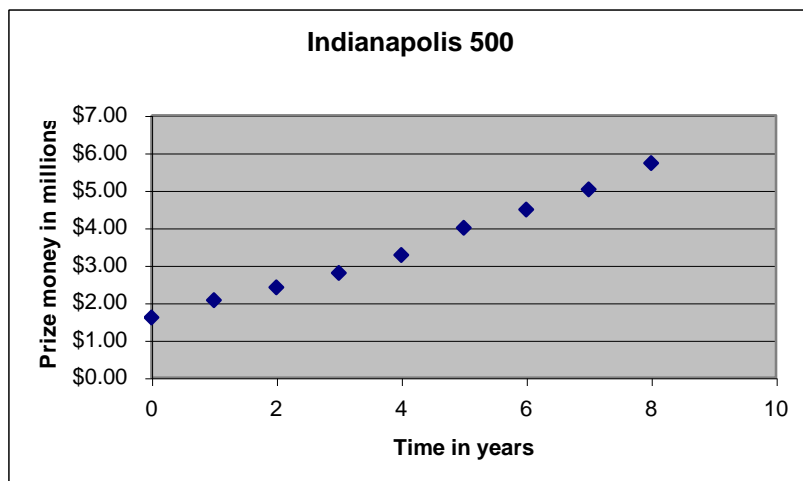
Example: The total prize money, in millions of dollars, awarded at the Indianapolis 500 race from 1981 to 1989 is given in the table below:

1. Table

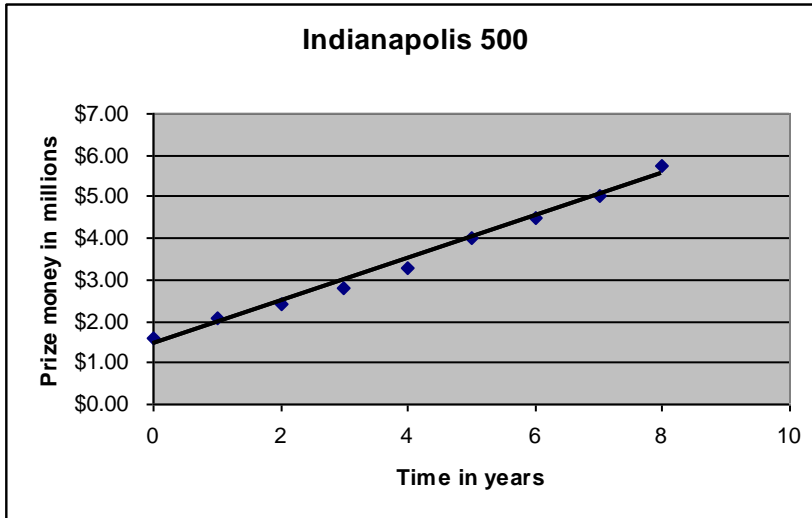| Year	Prize Money
1981	$1.605
1982	2.067
1983	2.411
1984	2.796
1985	3.271
1986	4.001
1987	4.490
1988	5.025
1989	5.724

2. Graph

Graph the data as a scatter plott:

3. Draw the line of best fit.
The line of best fit (LOBF) approximates the data trend by passing closest to all the points.

Indianapolis 500

4. Find the rate of change of the line.
Find the approximate rate of change of your line by choosing any two points on the line. For example you might choose (1, 2) and (7,5)

Divide the change in prize money by the change in years as indicated by the two points.

$$Rate\,of\,Change = \frac{Change\,in\,\Pr ize\,Money}{Change\,in\,Years} = \frac{5-2}{7-1} = \frac{3\,million\,dollars}{6\,years} = .5\,million\,dollars\,per\,year$$

The rate of change shows how much the amount of prize money is changing each year.

5. Find the equation of the line. (The function)
Find the approximate equation for the line using either of the points and the rate of change:

For example you might use the point (1, 2)
Let   P = Amount of prize money in millions of dollars
       t = time in years
The rate of change is .5
       P – 2 = .5 (t – 1) or
       P = .5t + 1
The equation means that to find the amount of prize money in millions in any year you must multiply the time since you started in 1981 by .5 and add 1.
For this equation P is a function of t; the value of P depends on t. P is the dependent variable and t is the independent variable for the function.

6. Check your model.

To check your model, use the equation to compare values given by the model to the actual data values. The closer your model values are to the actual data values, the more accurate your model is.

For 1985:   Actual value of the prize money is $3.271 million
                    Model value = .5(5) + 1 = $3.5 million

7. Use your model to predict the future

To predict the approximate future value for the prize money, use your equation for a future time:
   For 1995, the time is 14 years after 1981 so t = 14.
According to your model P = .5t +1
So P = .5(14) + 1
P = $8 Million dollars in 1995

8. Use your model to predict a past value

To find an approximate past value for the prize money, use your equation for a past time.

For 1980, the time is –2 years after 1981 so t = -1
According to your model P = .5t + 1
So P = .5(-1) +1
P = $.5 Million dollars in 1980

9. Use your model to predict the inverse of the function

To find the approximate time that a certain amount of prize money was awarded, transform the equation and evaluate it for the amount of time.

For $6 million dollars,
P = .5t+1
P – 1 = .5 t
t = (P-1)/.5
t = (6-1)/.5
t = 10 years after 1981
In 1991 the prize money will be approximately $6 million dollars.

10. Explain the limitations of the model

a. Accuracy: For 1985: Actual value of the prize money is $3.271 million
                    Model value = .5(5) + 1 = $3.5 million
                    The model point is within 7% of the actual value.
                    To find the percent error subtract the actual value from the model value
                    Then divide the result by the actual value and multiply by 100.

                    Excel can be used to make a scatter plot of the data, with an equation and
                    trendline. The r^2 value of the trendline is the correlation coefficient that

describes the accuracy of the model. A perfect model has a correlation coefficient of +1 or -1. The closer the correlation coefficient is to these values, the more accurate the model.

For this model the Excel gives an equation of

$y = 0.511x + 1.44$ or $P = .511t + 1.44$

$R^2 = 0.99$

The correlation coefficient of .99 is very close to 1 so this model is quite accurate.

b. Outliers: Outliers are points that are far outside the line of best fit. These may be due to data errors or to unusual circumstances. To find the average behavior, sometimes outliers must be omitted from the model.

c. Errors: Possible errors that could have occurred during data collection for this model.

11. Describe the model in words.

This model shows that starting with about 1.44 million dollars in 1981 in the amount of prize money for the Indianapolis 500 is increasing at the rate of about .5 million dollars each year. If the trend continues, the amount of prize money will be 8 million dollars in 1995.

Note: The actual amount of prize money in 1995 was 8 million dollars, so this prediction was correct.

The prediction for 2007 would be 14 million dollars but the actual amount was only 10.7 million dollars. The model is not valid for 2007.

12. Describe your data source.

Example:

Data Source: Johnson's Indy 500

http://www.geocities.com/johnsonindy500/indy500/indyprizemoney.html

Checklist for Models
☐ Data Table
☐ Graph
☐ Line of Best Fit
☐ Rate of Change
☐ Equation
☐ Test the Model
☐ Predict a Future Value
☐ Predict a Past Value
☐ Predict the Inverse of the Function
☐ Explain the model limitations
☐ Describe the model in words
☐ Describe the data source