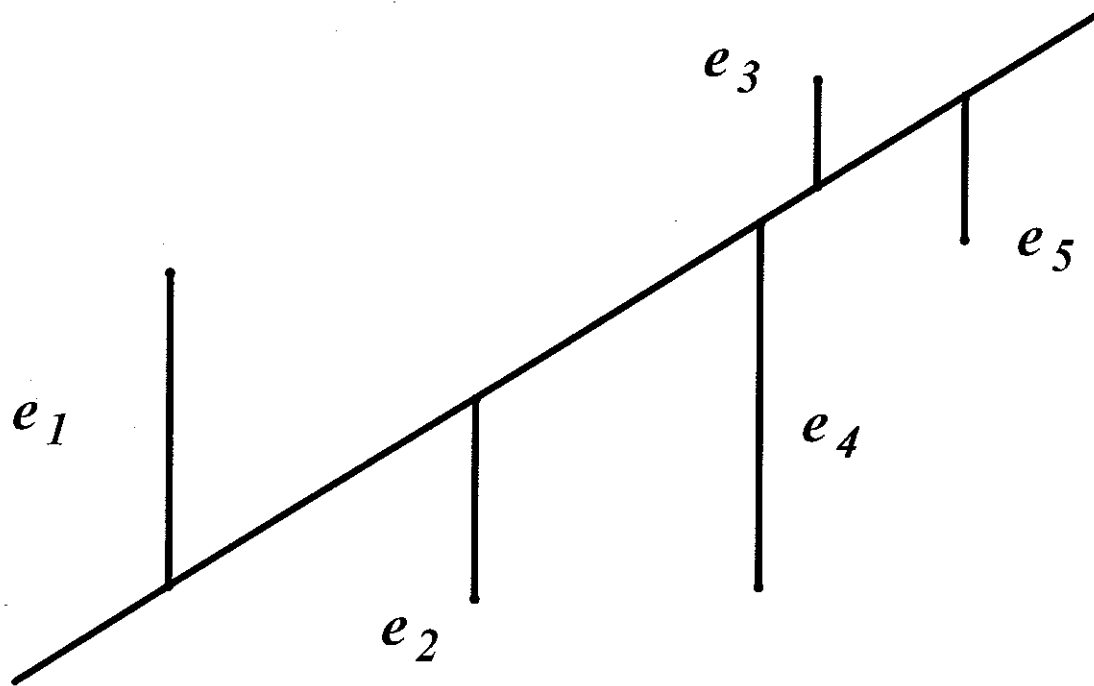


Regression: The Method of Least Squares



Choose the line that makes the sum of the squared errors as small as possible.

The regression line minimizes $(e_1)^2 + (e_2)^2 + (e_3)^2 + (e_4)^2 + (e_5)^2$.

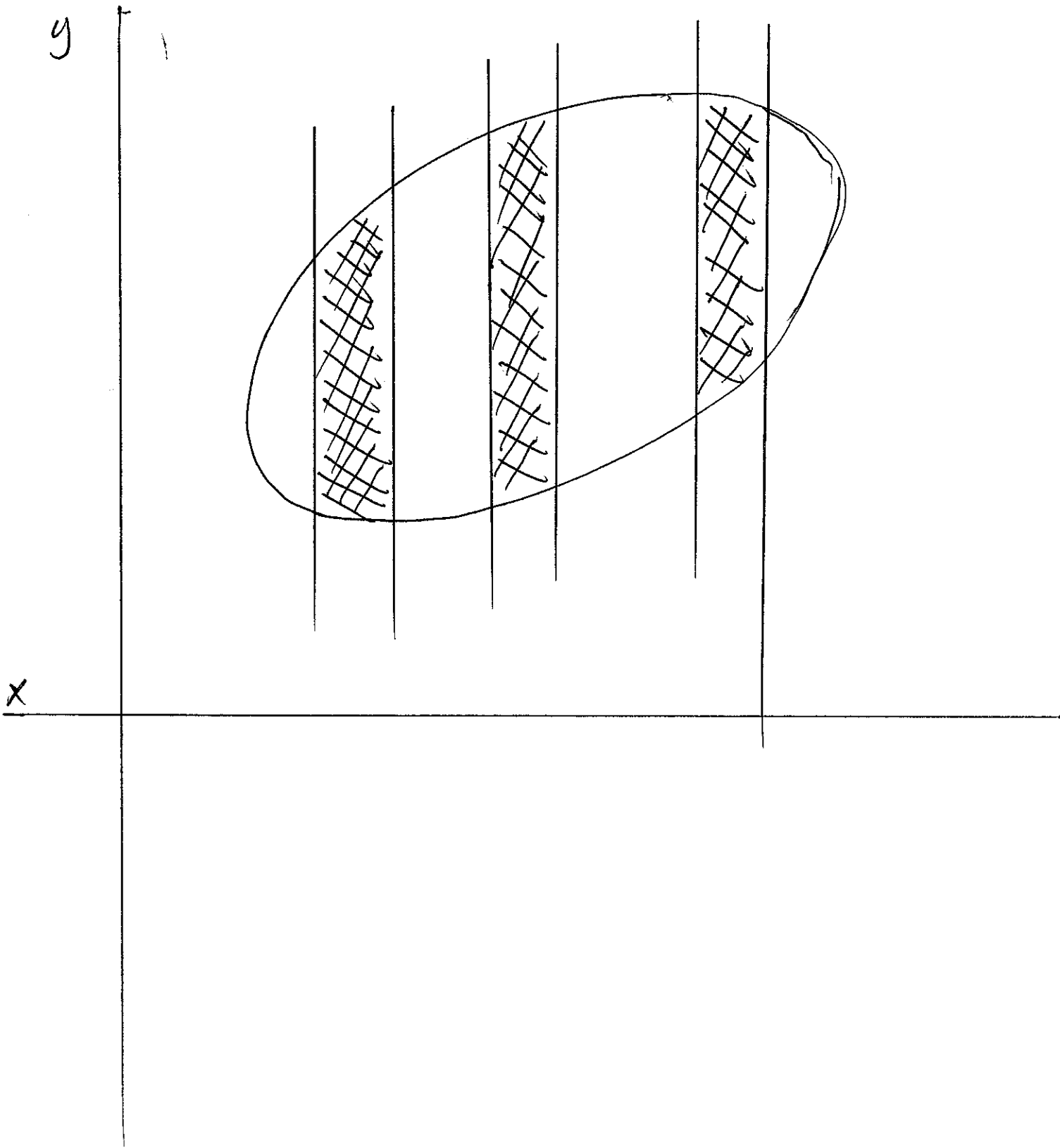
Among all lines, the regression line makes the smallest r.m.s. error in predicting y from x .

The r.m.s. error is the r.m.s. size of the errors.

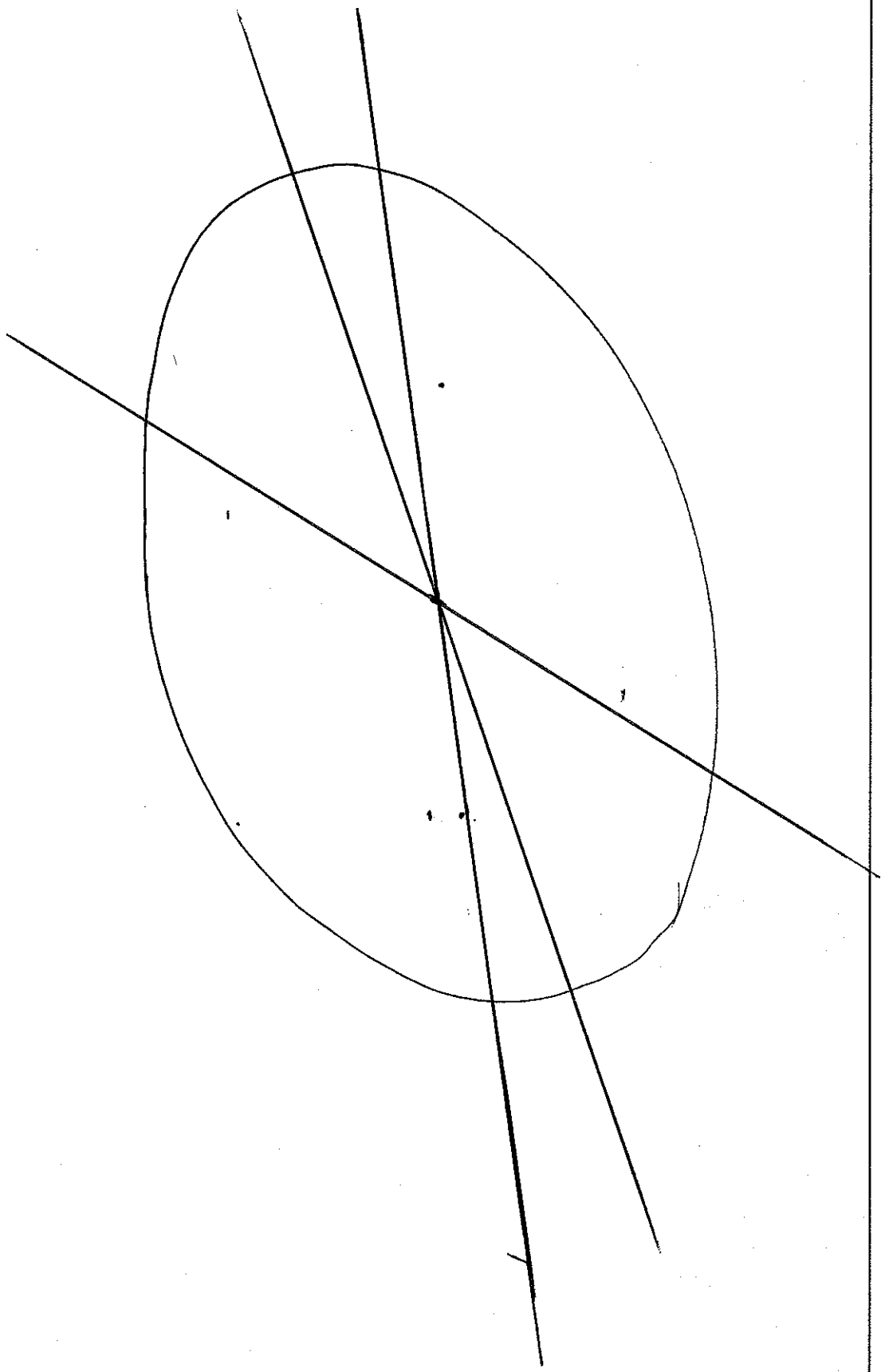
The r.m.s. error measures how good a prediction is. It says how large the errors are likely to be.

To calculate the r.m.s. error use the following shortcut:

$$\text{r.m.s. error} = \sqrt{(1 - r^2)} (SD_Y)$$

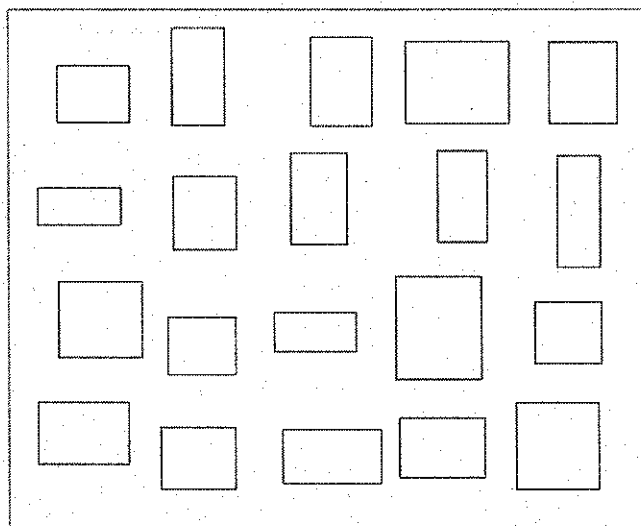


X



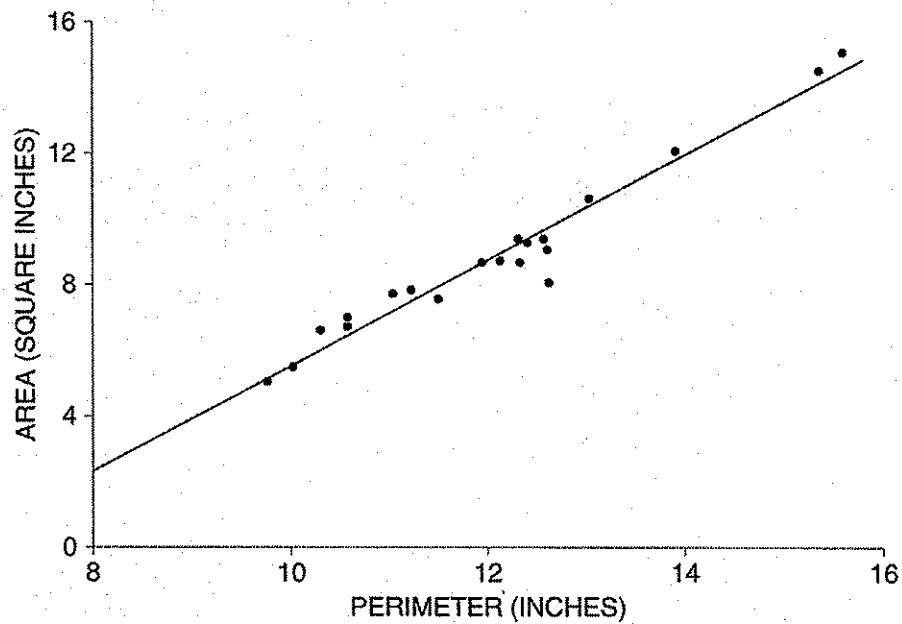
A bad regression example

Measure the area and perimeter of these rectangles

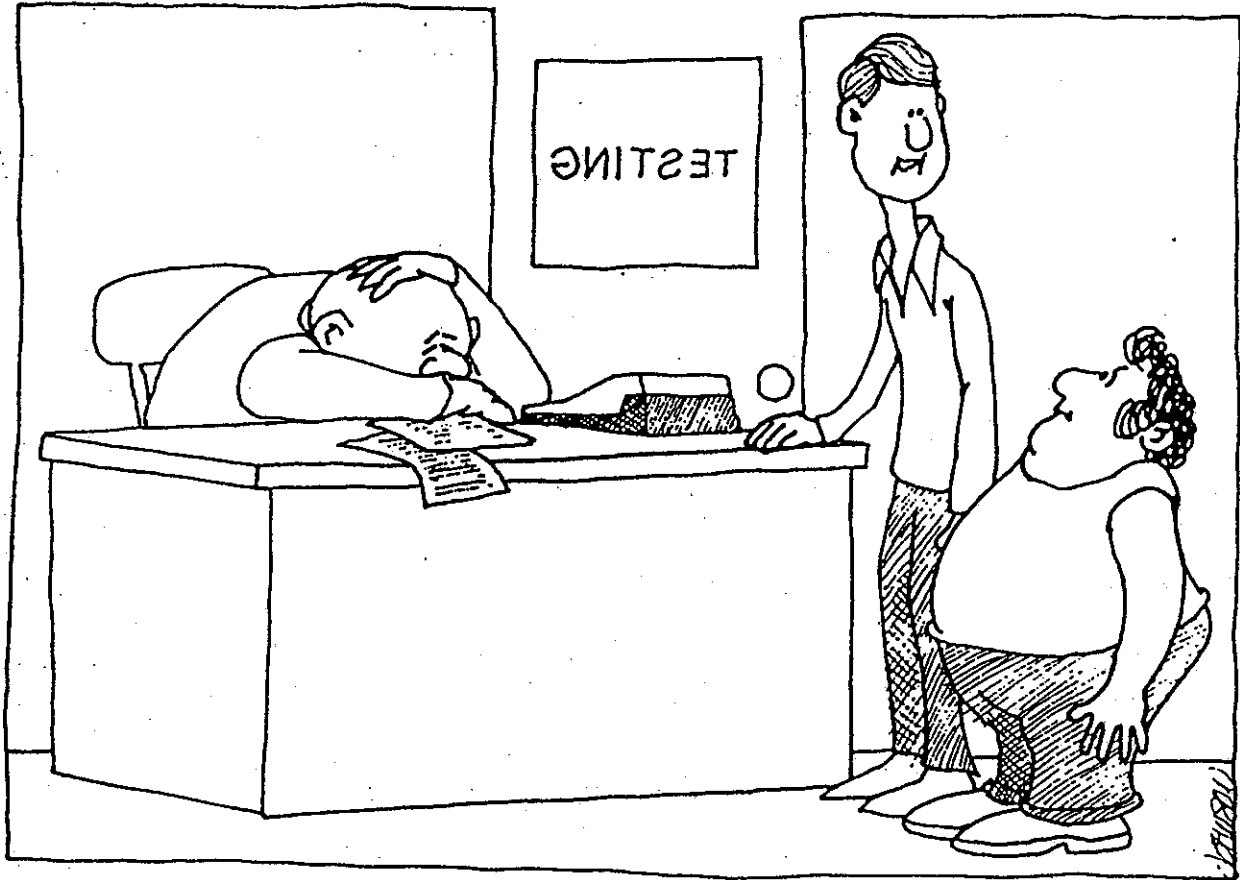


SCALE: 0 2 4 6 8 10 INCHES

The correlation between area and perimeter is $r = 0.98$! The scatter diagram:

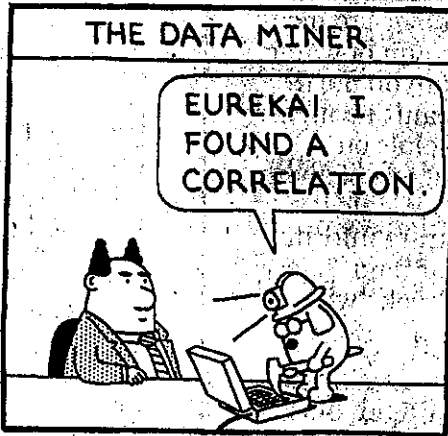


$$x = 1000, \quad y = \frac{1}{10^{10}}$$



"He says we've ruined his positive association between height and weight."

DILBERT



www.dilbert.com scottadams@aol.com



© 1999 United Feature Syndicate, Inc.

